TESIS-SS14 2501

MODEL REGRESI SEMIPARAMETRIK BIRESPON DENGAN PENDEKATAN DERET FOURIER

M FARIZ FADILLAH MARDIANTO
NRP 1313 201 050

DOSEN PEMBIMBING
Prof. Dr. Drs. I Nyoman Budiantara, M.Si

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

TESIS-SS14 2501

SEMIPARAMETRIC REGRESSION MODEL OF BIRESPONSE USING FOURIER SERIES

M. FARIZ FADILLAH MARDIANTO
NRP 1312 201 910

SUPERVISOR
Prof. Dr. Drs. I Nyoman Budiantara, M.Si

PROGRAM OF MAGISTER
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2014

MODEL REGRESI SEMIPARAMETRIK BIRESPON DENGAN PENDEKATAN DERET FOURIER

Tesis ini disusun untuk memenuhi salah satu syarat memperoleh gelar

Magister Sains (M.Si)

di

Institut Teknologi Sepuluh Nopember

Oleh:

M. FARIZ FADILLAH MARDIANTO

NRP. 1313 201 050

Tanggal Ujian : 16 Januari 2015

Periode Wisuda : Maret 2015

Disetujui Oleh :

1. Prof. Dr. Drs. I Nyoman Budiantara, M.Si (Pembimbing)
NIP. 19650603 198903 1 003

2. Dr. Vita Ratnasari, S.Si, M.Si (Penguji)
NIP. 19700910 199702 2 001

3. Dr. Kartika Fithriasari, M.Si (Penguji)
NIP. 19691212 199303 2 002

Direktur Program Pasca Sarjana ITS,

Prof. Dr. Ir. Adi Soeprijanto, M.T.
NIP. 19640405 199002 1 001

MODEL REGRESI SEMIPARAMETRIK BIRESPON DENGAN PENDEKATAN DERET FOURIER

Nama Mahasiswa : M.Fariz Fadillah Mardianto
NRP : 1313 201 050
Pembimbing : Prof. Dr. Drs I Nyoman Budiantara, M.Si

ABSTRAK

Dalam regresi, terdapat tiga pendekatan yaitu pendekatan parametrik, nonparametrik, dan semiparametrik. Pendekatan regresi semiparametrik terdiri atas dua komponen yaitu komponen parametrik dan nonparametrik. Dalam pemilihan ini, komponen parametrik yang dikembangkan berbentuk linier, sedangkan komponen nonparametrik yang dikembangkan menggunakan Deret Fourier. Diberikan model regresi semiparametrik birespon dengan variabel komponen parametrik $x_1, x_2 \ldots, x_p$, dan variabel komponen nonparametrik $t_1, t_2 \ldots, t_R$ mengikuti model :

$$y_{ji} = \beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi} + \sum_{r=1}^{R} f_j(t_{ri}) + \varepsilon_{ji}$$

dimana $i = 1,2,\ldots,n$ dan $j = 1,2$. Komponen nonparametrik didekati dengan fungsi Deret Fourier berbentuk :

$$f_j(t_{ri}) = \gamma_{jr} t_{ri} + \frac{\alpha_{0jr}}{2} + \sum_{k=1}^{K} (\alpha_{kjr} \cos k t_{ri})$$

Parameter K merupakan parameter osilasi. Berdasarkan optimasi WLS (*Weighted Least Square*) diperoleh estimator untuk komponen parametrik

$$\hat{\boldsymbol{\beta}} = \mathbf{M}\left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1}\left\{\mathbf{X}^T - \mathbf{X}^T\mathbf{W}\mathbf{T}\left(\mathbf{T}^T\mathbf{W}\mathbf{T}\right)^{-1}\mathbf{T}^T\right\}\mathbf{W}\mathbf{y} = \mathbf{A(K)y}$$

dan estimator untuk komponen nonparametrik

$$\hat{\boldsymbol{\eta}} = \mathbf{N}\left(\mathbf{T}^T\mathbf{W}\mathbf{T}\right)^{-1}\left\{\mathbf{T}^T - \mathbf{T}^T\mathbf{W}\mathbf{X}\left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^T\right\}\mathbf{W}\mathbf{y} = \mathbf{B(K)y}$$

Selanjutnya diperoleh estimator model regresi semiparametrik birespon dengan pendekatan Deret Fourier

$$\hat{\mathbf{y}} = (\mathbf{XA(K)} + \mathbf{TB(K)})\mathbf{y} = \mathbf{C(K)y}$$

Estimator tersebut digunakan untuk memodelkan Indeks Pembangunan Manusia (IPM) dan Indeks Mutu Hidup (IMH) di Jawa Timur. Dengan menggunakan metode GCV diperoleh model regresi semiparametrik birespon

$$\hat{y}_{1i} = 0{,}043 + 0{,}326x_{1i} + 0{,}354x_{2i} + 0{,}182x_{3i} + 0{,}337x_{4i} - 0{,}042t_{1i} + 0{,}053\cos t_{1i} - 0{,}119\cos 2t_{1i} - 0{,}146\cos 3t_{1i} + 0{,}08t_{2i} + 0{,}035\cos t_{2i} + 0{,}035\cos 2t_{2i} + 0{,}035\cos 3t_{2i}$$

$$\hat{y}_{2i} = -1{,}015 + 0{,}939x_{1i} + 0{,}373x_{2i} + 0{,}842x_{3i} - 0{,}31x_{4i} - 0{,}005t_{1i} + 0{,}226\cos t_{1i} - 0{,}636\cos 2t_{1i} - 1{,}314\cos 3t_{1i} + 0{,}525t_{2i} - 1{,}029\cos t_{2i} - 1{,}029\cos 2t_{2i} - 1{,}029\cos 3t_{2i}$$

dengan nilai GCV sebesar 753.586, MSE sebesar 3,052 dan R^2 91,98%. Model tersebut dapat digunakan sebagai alternatif dari model regresi parametrik birespon yang melanggar semua asumsi klasik residual.

Kata kunci: Regresi Semiparametrik, Birespon, Deret Fourier, *Weighted Least Square*, IPM, IMH.

"halaman ini sengaja dikosongkan"

SEMIPARAMETRIC REGRESSION MODEL OF BIRESPONSE USING FOURIER SERIES

Name	: M.Fariz Fadillah Mardianto
NRP	: 1313 201 050
Supervisor	: Prof. Dr. Drs I Nyoman Budiantara, M.Si

ABSTRCT

In regression analysis, there are three approaches such as parametric regression, nonparametric regression, and semiparametric regression approach. Semiparametric regression approach consists of two components, the parametric and nonparametric components. In this choice, the form of the parametric components that will be developed is a linear form, and the nonparametric components that will be developed using Fourier series. Consider a biresponse semiparametric regression model that consists component parametric variables $x_1, x_2 \ldots, x_p$, and component nonparametric variables $t_1, t_2 \ldots, t_R$. It follows a model:

$$y_{ji} = \beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi} + \sum_{r=1}^{R} f_j(t_{ri}) + \varepsilon_{ji}$$

$i = 1,2,\ldots,n$ and $j = 1,2$. Nonparametric component is approached with Fourier series:

$$f_j(t_{ri}) = \gamma_{jr} t_{ri} + \frac{\alpha_{0jr}}{2} + \sum_{k=1}^{K} (\alpha_{kjr} \cos k t_{ri})$$

K is oscillation parameter. Using WLS (*Weighted Least Square*) to optimize, can be derived an estimator for parametric component

$$\hat{\boldsymbol{\beta}} = \mathbf{M}\left(\mathbf{X}^{T}\mathbf{W}\mathbf{X}\right)^{-1}\left\{\mathbf{X}^{T} - \mathbf{X}^{T}\mathbf{W}\mathbf{T}\left(\mathbf{T}^{T}\mathbf{W}\mathbf{T}\right)^{-1}\mathbf{T}^{T}\right\}\mathbf{W}\mathbf{y} = \mathbf{A(K)y}$$

And an estimator for nonparametric component

$$\hat{\boldsymbol{\eta}} = \mathbf{N}\left(\mathbf{T}^{T}\mathbf{W}\mathbf{T}\right)^{-1}\left\{\mathbf{T}^{T} - \mathbf{T}^{T}\mathbf{W}\mathbf{X}\left(\mathbf{X}^{T}\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^{T}\right\}\mathbf{W}\mathbf{y} = \mathbf{B(K)y}$$

After that can be derived an estimator for biresponse semiparametric regression using Fourier series:

$$\hat{\mathbf{y}} = (\mathbf{XA(K)} + \mathbf{TB(K)})\mathbf{y} = \mathbf{C(K)y}$$

That estimator can be used to make a model for Human Development Index (HDI) and Physical Quality of Life Index (PQLI) in East Java. Using GCV, the model is:

$$\hat{y}_{1i} = 0{,}043 + 0{,}326 x_{1i} + 0{,}354 x_{2i} + 0{,}182 x_{3i} + 0{,}337 x_{4i} - 0{,}042 t_{1i} + 0{,}053 \cos t_{1i} - 0{,}119 \cos 2t_{1i} - 0{,}146 \cos 3t_{1i} + 0{,}08 t_{2i} + 0{,}035 \cos t_{2i} + 0{,}035 \cos 2t_{2i} + 0{,}035 \cos 3t_{2i}$$

$$\hat{y}_{2i} = -1{,}015 + 0{,}939 x_{1i} + 0{,}373 x_{2i} + 0{,}842 x_{3i} - 0{,}31 x_{4i} - 0{,}005 t_{1i} + 0{,}226 \cos t_{1i} - 0{,}636 \cos 2t_{1i} - 1{,}314 \cos 3t_{1i} + 0{,}525 t_{2i} - 1{,}029 \cos t_{2i} - 1{,}029 \cos 2t_{2i} - 1{,}029 \cos 3t_{2i}$$

GCV value 753.586, MSE 3,052 and R^2 91,98%. This model can be used as alternative model from biresponse parametric regression because against classical assumptions in regression analysis.

Keyword : Semiparametric Regression, Biresponse, Fourier Series, Weighted Least Square, HDI, PQLI

"halaman ini sengaja dikosongkan"

KATA PENGANTAR

Puji syukur dipanjatkan atas kehadirat Allah SWT karena atas segala limpahan rahmat, dan hidayah yang dicurahkan serta segala petunjuk dan kemudahan yang diberikan, Alhamdulillah penulis dapat menyelesaikan laporan tesis dengan judul "**Model Regresi Semiparametrik Birespon dengan Pendekatan Deret Fourier**".

Laporan Tesis ini dapat selesai tak lepas dari peranan, dukungan dan motivasi berbagai pihak. Oleh karena itu pada kesempatan ini penulis ingin mengucapkan terima kasih yang sedalam-dalamnya kepada:

1. Kedua orang tua yang sangat saya cintai dan hormati, Bapak M. Taufan dan Ibu Sunarsih, beserta adik tersayang M. Rama Fadillah Dessyanto terima kasih atas segala doa, dukungan, dan motivasi yang tiada henti. Semoga bapak, ibu, dan adik senantiasa diberi kesehatan, rezeki, dan kelancaran dalam menjalani aktivitas.
2. Pihak Pasca Sarjana ITS yang telah memberikan kesempatan untuk melanjutkan studi melalui program beasiswa Fresh Graduate ITS.
3. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen pembimbing yang telah memberikan banyak masukan dan motivasi serta dengan sabar mengarahkan mahasiswa bimbingannya menyelesaikan Tesis ini.
4. Ibu Dr. Vita Ratnasari, S.Si, M.Si dan Ibu Dr. Kartika Fithriasari, M.Si selaku dosen penguji yang telah memberikan banyak saran, kritik, serta masukan demi kesempurnaan Tesis ini.
5. Bapak Dr. Muhammad Mashuri, M.T, selaku ketua Jurusan Statistika FMIPA ITS, terima kasih banyak atas segala kemudahan urusan akademis dan fasilitas yang menunjang di Jurusan Statistika ini.
6. Bapak Dr. Suhartono, M.Sc selaku kaprodi pasca sarjana Jurusan Statistika-ITS, terima kasih atas motivasinya.

7. Bapak dan Ibu dosen pengajar khusuusnya di Jurusan Statistika dan Matematika FMIPA ITS, terima kasih atas ilmu yang telah diberikan.
8. Panitia Semnas Matematika 2014 di Universitas Udayana Denpasar Bali, terima kasih telah memberikan waktu memaparkan sebagian isi dari Tesis ini.
9. Segenap pihak Badan Pusat Statistik (BPS) Provinsi Jawa Timur yang telah banyak membantu penyediaan data.
10. Oka dan Evelin, terima kasih atas pencerahan pemrograman tesis ini.
11. Buat orang yang paling spesial, terima kasih motivasi dan dukungan spesialnya.
12. Rekan – rekan seperjuangan sebimbingan, Mirah, Nanik, Bobby, dan Bu Lilis yang memberikan semangat, masukan, dan terima kasih telah menjadi teman diskusi dan tukar ide.
13. Rekan – rekan S2 Statistika, Mas Amin, Reza, Ditago, Iqbal, Adi, Mbak Fenda, Fifi, Tiwi, Kiki, Puspita, Fitri, Junta, Irma, Dea, Novri, Merly, Rizky, Ulyah dan teman-teman lainnya, semoga komunikasi yang baik tetap terjalin.
14. Teman-teman alumni Matematika ITS angkatan 2009, Indana, Ruzika, Himma, Ika, Lia, Izzah, Nining, Angga, Ali, Mico, Fahim, dan Kresna, terima kasih dukungannya.
15. Rekan – rekan di Data Analyst ADM, Bapak – Ibu dosen di Politeknik Sakti Surabaya, Fakultas Teknik Ubaya, terima kasih motivasinya.
16. Semua saudara dari bapak maupun ibu, terima kasih doa dan dukungannya.
17. Serta semua pihak yang tidak dapat saya sebutkan satu-persatu.

Semoga Tesis ini dapat bermanfaat dan dapat menambah wawasan keilmuan. Penulis menyadari bahwa Tesis ini masih membutuhkan saran dan kritik yang membangun dari semua pihak untuk perbaikan.

Surabaya, Januari 2015

Penulis

DAFTAR ISI

DAFTAR GAMBAR

"halaman ini sengaja dikosongkan"

DAFTAR TABEL

"halaman ini sengaja dikosongkan"

DAFTAR LAMPIRAN

“halaman ini sengaja dikosongkan”

BAB 1
PENDAHULUAN

1.1. Latar Belakang

Metode regresi adalah suatu metode dalam statistika yang digunakan untuk mengetahui pola hubungan antara variabel prediktor x_i dan variabel respon y_i. Diberikan data berpasangan antara variabel prediktor dan respon (x_i, y_i) yang mengikuti model regresi

$$y_i = f(x_i) + \varepsilon_i \ , i = 1,2, \dots, n \tag{1.1}$$

dengan n yaitu banyaknya pengamatan. Fungsi $f(x_i)$ disebut fungsi regresi, dan ε_i adalah error random yang diasumsikan bersifat identik, independen, dan berdistribusi normal dengan mean nol dan variansi σ^2 (Eubank, 1988). Dalam analisis regresi, jika pola data diketahui misalnya berpola linier, kuadratik, atau kubik, maka metode regresi yang digunakan adalah metode regresi parametrik (Hardle, 1990; Islamiyati dan Budiantara, 2007). Namun dalam kenyataannya tidak semua data mengikuti pola – pola tertentu. Jika hubungan variabel prediktor dan respon tidak diketahui polanya, maka metode regresi nonparametrik sesuai untuk memodelkan hubungan antar variabel tersebut (Budiantara *et al*., 2010).

Selain regresi parametrik dan regresi nonparametrik, terdapat juga regresi semiparametrik yang merupakan kombinasi antara regresi parametrik dan regresi nonparametrik. Kombinasi ini dimaksudkan bahwa dalam regresi semiparametrik memuat komponen parametrik dan komponen nonparametrik. Regresi semiparametrik ini muncul karena adanya kasus-kasus pemodelan dimana hubungan antar variabelnya selain ada yang linier, juga ada yang tidak diketahui polanya. Dalam penelitian ini regresi semiparametrik yang digunakan adalah kombinasi komponen parametrik linier dan komponen nonparametrik dengan pendekatan Deret Fourier.

Model regresi nonparametrik dengan pendekatan Deret Fourier (Antoniadis *et al*, 1994) merupakan salah satu model dalam regresi nonparametrik. Adapun model–model regresi nonprametrik lainnya adalah Spline (Budiantara, 2002 ; Budiantara *et al.,* 1997; Green dan Silverman, 1994; Wahba, 1990), Kernel

(Hardle, 1990), Wavelet (Antoniadis *et al*, 2001), dan sebagainya. Salah satu keunggulan pendekatan regresi nonparametrik dengan menggunakan Deret Fourier adalah mampu mengatasi data yang mempunyai pola berulang, yaitu pengulangan terhadap nilai variabel respon untuk variabel prediktor yang berbeda-beda (Pane *et al*., 2014).

Penelitian menggunakan Deret Fourier dalam regresi nonparametrik maupun semiparametrik telah banyak dikembangkan. Penelitian mengenai Deret Fourier antara lain oleh Bilodeau (1992), dan Tripena (2007) yang mengkaji estimator Deret Fourier pada regresi nonparametrik. Semiati (2010) mengembangkan estimasi model regresi nonparametrik Deret Fourier birespon. Prahutma (2013) menerapkan pendekatan regresi nonparametrik Deret Fourier pada permasalahan pengangguran terbuka di Jawa Timur. Bidermann, Datte, dan Hoffmann (2009) dalam jurnalnya meneliti rancangan optimal untuk mendapatkan suatu *constrained* dalam model regresi Fourier. Regresi semiparametrik satu respon menggunakan Deret Fourier telah dikembangkan oleh Asrini (2011). Asrini (2014) menerapkan penelitiannya pada data produksi padi di Jawa Tengah.

Penelitian–penelitian ini belum ada yang meneliti pendekatan Deret Fourier dalam regresi semiparametrik untuk kasus birespon. Dalam kenyataannya banyak aplikasi yang dapat dimodelkan oleh regresi semiparametrik dengan pendekatan Deret Fourier dimana variabel responnya lebih dari satu.

Data yang digunakan dalam penelitian ini adalah IPM (Indeks Pembangunan Manusia) yang disebut juga *Human Development Index* dan IMH (Indeks Mutu Hidup) yang disebut juga *Phsyical Quality of Life Index* dimana berkaitan dengan indikator kesejahteraan daerah atau capaian pembangunan suatu daerah (Badan Pusat Statistik, 2008). Beberapa capaian pembangunan manusia dalam lingkup daerah dapat dinilai berdasarkan indikator kesejahteraan yang ada dengan komponen penyusun tiap indikator kesejahteraan daerah yang berbeda. Persoalannya adalah capaian pembangunan manusia sangat bervariasi dimana beberapa aspek pembangunan tertentu berhasil dan beberapa aspek pembangunan lainnya belum berhasil (UNDP,2013).

Dewasa ini persoalan mengenai capaian pembangunan manusia telah menjadi perhatian para penyelenggara pemerintahan. Berbagai ukuran pembangunan manusia dibuat namun tidak semuanya dapat digunakan sebagai ukuran standar yang dapat dibandingkan antar wilayah. Oleh karena itu Badan Perserikatan Bangsa-Bangsa (PBB) menetapkan suatu ukuran standar pembangunan manusia yaitu indeks pembangunan manusia (IPM). Indeks ini dibentuk berdasarkan indikator–indikator yaitu angka harapan hidup, angka melek huruf, rata-rata lama sekolah dan kemampuan daya beli. Indikator angka harapan hidup merepresentasikan dimensi panjang umur dan sehat. Selanjutnya, angka melek huruf dan rata-rata lama sekolah mencerminkan output dari dimensi pengetahuan. Adapun indikator kemampuan daya beli digunakan untuk mengukur dimensi hidup layak. Indikator pendidikan dan kesehatan juga mempengaruhi capaian pembangunan manusia. Luasnya cakupan pembangunan manusia menjadikan peningkatan IPM sebagai manifestasi dari pembangunan manusia yang dapat ditafsirkan sebagai keberhasilan dalam meningkatkan kemampuan dalam memperluas pilihan-pilihan (*enlarging the choices of the people*) untuk kesejahteraan (UNDP, 2013).

Beberapa penelitian tentang IPM di Jawa Timur telah dilakukan tanpa melibatkan peran indikator kesejahteraan lain seperti IMH, diantaranya Perdana (2012) yang meneliti hubungan variabel–variabel IPM di Jawa Timur dengan tingkat kemiskinan, Melliana (2013) yang menggunakan regresi panel untuk menganalisis faktor yang mempengaruhi IPM kabupaten dan kota di Jawa Timur, dan sebagainya. Penelitian–penelitian tersebut belum ada yang mengkaji tentang indeks kesejahteraan daerah dan pembangunan lain seperti IMH (Indeks Mutu Hidup).

Untuk mengetahui capaian kesejahteraan daerah dan pembangunan suatu daerah juga dapat dilihat dengan IMH. BPS hanya menggunakan IPM sebagai alat ukurnya, namun di negara lain IMH juga dijadikan acuan melihat capaian pembangunan suatu daerah. David Morris (1970) mengembangkan IMH untuk *Overseas Development Council*, sebagai salah satu dari sejumlah solusi atas ketidakpuasan menggunakan GNP sebagai indikator pembangunan (Roger Nosal dan Tom Schultz, 2008). IMH menarik untuk diteliti lebih lanjut karena indikator ini sama dengan IPM dimana terbentuk berdasarkan indikator–indikator di bidang pendidikan dan kesehatan. IPM dan IMH penting untuk mengetahui keberhasilan pembangunan suatu daerah, mengetahui kualitas penduduk di suatu daerah, dan evaluasi kebijakan pembangunan di suatu daerah.

Dalam penelitian ini dimodelkan IPM dan IMH untuk Provinsi Jawa Timur. Provinsi yang beribukota di Surabaya ini memiliki perkembangan IPM yang terus meningkat dari tahun ke tahun (Badan Pusat Statistik, 2013). Kondisi IPM di Jawa Timur menarik diteliti karena IPM provinsi yang terdiri dari 38 kabupaten dan kota ini terus mengalami peningkatan bahkan hampir mendekati IPM Indonesia di tahun 2012. Dari data BPS IPM Indonesia tahun 2011 sebesar 72.77 dan Jawa Timur sebesar 72,18 dan berada di urutan 17, sedangkan tahun 2012, IPM Indonesia sebesar 73,29 dan Jawa Timur sebesar 72,83 berada di urutan 15 (Badan Pusat Statistik, 2013).

Berdasarkan hal tersebut, dalam penelitian ini dikaji bentuk estimator regresi semiparametrik birespon dengan pendekatan Deret Fourier. Selanjutnya penelitian ini memodelkan IPM dan IMH di Jawa Timur menggunakan regresi semiparametrik birespon dengan pendekatan Deret Fourier.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, permasalahan yang dapat dirumuskan dalam penelitian ini adalah:

1. Bagaimana estimasi model regresi semiparametrik Deret Fourier birespon.
2. Bagaimana pola hubungan antara IPM dan IMH dengan variabel Indeks Kesehatan Masyarakat, Indeks Pendidikan, PDRB (Prooduk Domestik Regional Bruto), Indeks PPP (*Purchasing Power Parity*), TPAK (Tingkat Partisipasi Angkatan Kerja), dan Persentase Kemiskinan di Jawa Timur.

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Mendapatkan estimasi model regresi semiparametrik Deret Fourier birespon.
2. Memodelkan hubungan antara IPM dan IMH di Jawa Timur dengan variabel – variabel yang mempengaruhinya untuk perencanaan kebijakan kedepan.

1.4. Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah:

1. Memberikan kontribusi keilmuan khususnya pengembangan ilmu, baik dalam bidang statistika dalam hal ini metode regresi dan matematika yaitu Deret Fourier yang dapat diterapkan pada persoalan sosial dan kemasyarakatan.
2. Hasil dari penelitian ini dapat dijadikan bahan rujukan untuk mengambil kebijakan, dalam hal ini perencanaan target IPM dan IMH kedepannya di wilayah Jawa Timur.

1.5. Batasan Masalah

Dalam penelitian ini hanya dibatasi oleh hal – hal berikut :

1. Estimator regresi semiparametrik Deret Fourier diperoleh dengan metode WLS (*Weighted Least Square*)
2. Pemilihan parameter osilasi optimal menggunakan metode GCV (*Generalized Cross Validation*).
3. Parameter osilasi atau K yang digunakan dalam metode GCV dibatasi pada K=1, 2, dan 3.
4. Untuk menentukan variabel prediktor parametrik dan nonparametrik dilihat berdasarkan *scatter plot* dan uji linieritas.

"halaman ini sengaja dikosongkan"

BAB 2

TINJAUAN PUSTAKA

2.1. Analisis Regresi

Analisis regresi adalah suatu analisis statistika yang digunakan untuk menyelidiki hubungan antara variabel prediktor dengan variabel respon (Drapper dan Smith, 1992). Persamaan (1.1) merupakan persamaan umum dalam model regresi. Dalam prakteknya, terdapat tiga pendekatan untuk mengestimasi kurva regresi, yaitu pendekatan regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Apabila diasumsikan bentuk kurva f diketahui, maka dapat digunakan pendekatan regresi parametrik. Namun apabila diasumsikan bahwa bentuk kurva f tidak diketahui, maka pendekatan yang digunakan adalah pendekatan regresi nonparametrik. Selain itu, apabila diasumsikan bahwa bentuk kurva f sebagian diketahui dan sebagian tidak diketahui, maka digunakan pendekatan regresi semiparametrik (Eubank, 1999; Budiantara, 2001).

2.2. Regresi Parametrik

Regresi parametrik merupakan metode statistik yang digunakan untuk mengetahui pola hubungan antara variabel prediktor dengan variabel respon, dengan asumsi bahwa bentuk pola hubungannya diketahui. Hubungan antara variabel respon dan variabel prediktor dalam model regresi dapat berpola linier dan dapat pula berpola nonlinier dalam parameter (Draper dan Smith, 1992).

Diberikan suatu model regresi parametrik linier multivariabel atau regresi linier berganda berikut :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \varepsilon_i, \text{ untuk } i = 1,2,\cdots,n \tag{2.1}$$

dengan y_i sebagai variabel respon, $x_{1i}, x_{2i}, \dots, x_{pi}$ sebagai variabel prediktor, $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ merupakan parameter yang tidak diketahui dan ε_i adalah *error* random yang independen berdistribusi normal dengan mean nol dan variansi σ^2. Jika dinyatakan dalam bentuk matriks, maka model regresi parametrik (2.1) menjadi

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.2}$$

dengan

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ 1 & x_{13} & x_{23} & \dots & x_{p3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

dan $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

Estimator untuk parameter-parameter model dapat diperoleh dengan menggunakan berbagai metode yang telah dikenal dalam statistika, yaitu *Least Square* atau *Maximum Likelihood Estimator* (Wahba, 1990). Estimator untuk parameter $\boldsymbol{\beta}$ diberikan oleh

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{2.3}$$

2.3. Regresi Parametrik Birespon dan Beberapa Pengujian yang Mendukung

Model regresi parametrik birespon dilakukan sebelum melakukan pemodelan regresi semiparametrik birespon untuk mengetahui kelemahan metode tersebut jika diaplikasikan untuk memodelkan IPM dan IMH. Model ini termasuk dalam model regresi multivariat karena terdiri dari lebih dari satu variabel respon yang saling berkorelasi dan satu atau lebih variabel variabel prediktor (Johnson dan Wichern, 2007). Jika terdapat variabel respon berjumlah 2 dan p variabel prediktor yaitu, maka model regresi multivariat untuk pengamatan ke- i adalah

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \beta_{21}x_{2i} + \dots + \beta_{p1}x_{pi} + \varepsilon_{1i}$$

$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \beta_{22}x_{2i} + \dots + \beta_{p2}x_{pi} + \varepsilon_{2i}$$

dengan i =1,2,…,n . Model ini dapat dibentuk menjadi matriks pada persamaan (2.2) dengan ukuran yang berbeda, dimana $E(\boldsymbol{\varepsilon}_{(q)}) = 0$ dan $cov(\boldsymbol{\varepsilon}_{(q)}, \boldsymbol{\varepsilon}_{(q)}) = \boldsymbol{\sigma}_{qq}\mathbf{I}$. Berdasarkan metode OLS diperoleh estimator untuk $\boldsymbol{\beta}$ seperti pada persamaan (2.3). Beberapa hal yang berhubungan dengan regresi ini adalah pengujian signifikansi dan pengujian asumsi.

2.3.1. Pengujian Signifikansi Serentak

Uji serentak adalah pengujian secara bersama-sama terhadap parameter yang terdapat di dalam persamaan regresi. *Uji Wilks Lambda* digunakan untuk pengujian serentak.

Hipotesis :

$H_0 : \beta_{11} = \beta_{12} \ldots = \beta_{62} = 0$ (model tidak signifikan)

H_1: paling sedikit ada satu $\beta_{pq} \neq 0$

dimana $p = 1,2,3,4,5,6$ dan $q = 1,2$

(model signifikan)

Statistik Uji :

$$\Lambda_{hitung} = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|} = \frac{\left|\mathbf{Y}^{\mathbf{T}}\mathbf{Y} - \widehat{\boldsymbol{\beta}}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{Y}\right|}{|\mathbf{Y}^{\mathbf{T}}\mathbf{Y} - n\bar{\boldsymbol{y}}\,\bar{\boldsymbol{y}}^{T}|} \qquad (2.4)$$

dengan :

Λ : nilai *Wilk's Lambda*

E : matriks *sum of square error*

H : matriks *sum of square treatment*

$\bar{\boldsymbol{y}}$: vektor rata–rata dari matriks **Y**

$\widehat{\boldsymbol{\beta}}^{\mathbf{T}}$: transpose matriks $\widehat{\boldsymbol{\beta}}$, matriks $\widehat{\boldsymbol{\beta}}$ ditentukan oleh

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_{01} & \hat{\beta}_{02} & \cdots & \hat{\beta}_{0q} \\ \hat{\beta}_{11} & \hat{\beta}_{12} & \cdots & \hat{\beta}_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\beta}_{p1} & \hat{\beta}_{p2} & \cdots & \hat{\beta}_{pq} \end{bmatrix}$$

$\Lambda_{tabel} = \Lambda_{\alpha,df,dg,dh} = \Lambda_{\alpha,q,p,n-p-1}$ adalah nilai tabel kritis untuk *Wilk's Lambda* dimana α merupakan tingkat signifikansi, dan $df = p, dg = q$ dan $dh = n-p-1$ adalah derajat kebebasan untuk *Wilks Lambda*.

Daerah Penolakan :

- Jika $\Lambda_{hitung} \leq \Lambda_{tabel}$ maka H_0 ditolak dimana secara keseluruhan estimator dari parameter tidak sama dengan nol sehingga model signifikan (Rencher, 2002)

2.3.2. Pengujian Signifikansi Individu

Pengujian ini bertujuan untuk melihat pengaruh setiap variabel prediktor terhadap variabel-variabel respon.

Hipotesis :

$H_0 : \beta_{lj} = 0$ (parameter model prediktor ke– l ; $l = 1,2,3,4,5,6$ terhadap respon ke– j ; $j = 1,2$ tidak berpengaruh signifikan)

$H_1 : \beta_{lj} \neq 0$ (parameter model prediktor ke– ; $l = 1,2,3,4,5,6$, terhadap respon ke– j ; $j = 1,2$ berpengaruh signifikan)

Statistik Uji :

$$T = \frac{\widehat{\boldsymbol{\beta}}_{\boldsymbol{lj}}}{SE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{lj}})} \tag{2.5}$$

Daerah Penolakan :

Jika $|t_{Hitung}| > t_{Tabel}$ untuk semua parameter β, maka H_0 ditolak, artinya parameter model prediktor ke– l ; $l = 1,2,3,4,5,6$, terhadap respon ke– j ; $j = 1,2$ berpengaruh signifikan (Rencher, 2002).

2.3.3. Pengujian Asumsi Residual Identik

Untuk menguji apakah residual data identik, maka asumsi yang harus dipenuhi dalam melakukan pemodelan regresi multivariat adalah residual memiliki matriks varian-kovarian ($\boldsymbol{\Sigma}$) yang homogen. Untuk menguji syarat ini dapat dipergunakan statistik uji *Box's M*.

Hipotesis :

H_0 : $\boldsymbol{\Sigma_1} = \boldsymbol{\Sigma_2} = \cdots = \boldsymbol{\Sigma_g} = \boldsymbol{\Sigma}$ (residual identik)

H_1 : minimal ada satu $\boldsymbol{\Sigma_i} \neq \boldsymbol{\Sigma_j}$ untuk $i \neq j, i = 1,2, \dots, g, j = 1,2, \dots, g$ (residual tidak identik)

Statistik Uji :

$$u = -2(1 - c_1) \ln M \tag{2.6}$$

dengan

$$\ln M = \tfrac{1}{2}\textstyle\sum_{i=1}^{g} v_i \ln|\boldsymbol{S}_i| - \tfrac{1}{2}(\textstyle\sum_{i=1}^{k} v_i) \ln|\boldsymbol{S}_{pool}|$$

$$\boldsymbol{S}_{pool} = \frac{\sum_{i=1}^{g} v_i S_i}{\sum_{i=1}^{k} v_i}$$

$$c_1 = \left[\sum_{i=1}^{g} \frac{1}{v_i} - \frac{1}{\sum_{i=1}^{k} v_i}\right]\left[\frac{2k^2 + 3k - 1}{6(k+1)(g-1)}\right]$$

$$v_i = n_i - 1$$

dimana

g : banyak kelompok

k : banyak variabel residual

$\mathbf{S_i}$: matriks varian-kovarian kelompok ke– i

n_i : jumlah observasi kelompok ke– i

u : nilai *Box's M*

Daerah Penolakan :

Jika $u > \chi^2_{tabel} = \chi^2_{\alpha,\frac{1}{2}(g-1)k(k+1)}$ maka H_0 ditolak yang berarti matriks varian-kovarian residual adalah heterogen dan dapat disimpulkan residual tidak identik atau p-value < α=0.05 (Rencher, 2002).

2.3.4. Pengujian Asumsi Residual Independen

Untuk menguji apakah residual data independen atau tidak digunakan uji *Bartlett Sphericity*.

Hipotesis :

H_0 : $\mathbf{R_e}$=$\mathbf{I}$ (residual independen)

H_1 : $\mathbf{R_e}$≠$\mathbf{I}$ (residual tidak independen)

Statistik Uji :

$$\chi^2_{hitung} = -\left\{n - 1 - \frac{2q+5}{6}\right\} ln|\mathbf{R_e}| \qquad (2.7)$$

dimana $\mathbf{R_e}$ adalah matriks korelasi residual berukuran $q \times q$. $\mathbf{I}$ adalah matriks identitas orde q. Teori lebih lanjut tentang pengujian ini ada di subbab 2.8.

Daerah Penolakan :

Jika $\chi^2_{hitung} > \chi^2_{tabel} = \chi^2_{\alpha,\frac{1}{2}q(q-1)}$ maka tolak H_0 sehingga residual tidak independen.(Morrison, 2005).

2.3.5. Pengujian Asumsi Residual Berdistribusi Normal Multivariat

Pengujian normal multivariat dilakukan untuk mengetahui bahwa residual berdistribusi normal multivariat. Diberikan $\boldsymbol{y} = [y_1, y_2, \dots, y_p]$ dengan p variabel dikatakan berdistribusi multivariat normal dengan rata-rata $\boldsymbol{\mu}$ dan variansi $\sum$ jika memiliki fungsi densitas sebagai berikut:

$$f(y_1, y_2, \dots, y_p) = \frac{1}{(2\pi)^{p/2}|\sum|^{p/2}} exp\left(\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^{\boldsymbol{T}}\sum^{-\boldsymbol{1}}(\boldsymbol{y}-\boldsymbol{\mu})\right)$$

Pemeriksaan multivariat normal dilakukan dengan cara membuat q-q plot dari ${d_i}^2$ dan q_i. Tahapan-tahapan dalam membuat q-q plot adalah :

a. Menentukan nilai vektor rata-rata $\bar{y}$ dan invers dari matriks varians kovarian.
b. Menentukan nilai ${d_i}^2$ yang merupakan jarak Mahalanobis setiap pengamatan dengan vektor rata-ratanya.
c. Urutkan ${d_i}^2$, ${d_{(1)}}^2 < {d_{(2)}}^2 < \cdots < {d_{(n)}}^2$.
d. Menentukan nilai $p_i = \frac{i-1/2}{n}, i = 1,2, \dots, n$
e. Menentukan nilai q_i yang didekati dengan $y_p^2\big((n-j-1/2)/n\big)$.
f. Membuat *scatter plot* dengan ordinat ${d_i}^2$ dan axis q_i.
g. Jika terdapat plot di sekitar 50%, nilai ${d_i}^2 < X^2_{p,50\%}$ dan membentuk pola garis lurus, maka dikatakan bahwa data berdistribusi multivariat normal.

Pengujian hipotesisnya diberikan sebagai berikut:

Hipotesis :

H_0 : Residual berdistribusi normal multivariat

H_1 : Residual tidak berdistribusi normal multivariat

Statistik Uji :

$$d_i^2 = (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \mathbf{S}^{-1} (\boldsymbol{y}_i - \bar{\boldsymbol{y}}) \quad (2.8)$$

Daerah Penolakan :

Jika nilai t tidak berada di sekitar 50 % diputuskan untuk tolak H_0, dimana t adalah banyaknya $d_i^2 \leq \chi^2_{tabel} = \chi^2_{q,\frac{1}{2}}$ (Johnson dan Wichern, 2007).

2.4. Regresi Nonparametrik

Pada kondisi tertentu, tidak selamanya kurva regresi diketahui bentuknya. Apabila diasumsikan bentuk kurva regresi tidak diketahui, maka disarankan untuk menggunakan pendekatan regresi nonparametrik.

Eubank (1999) memberikan model regresi nonparametrik berbentuk:

$$y_i = f(t_i) + \varepsilon_i,\ i = 1,2,\cdots,n \tag{2.9}$$

dengan y_i adalah variabel respon. Fungsi $f(t_i)$ tidak diketahui bentuknya dan t_i merupakan variabel prediktor, serta ε_i diasumsikan berdistribusi $N(0,\sigma^2)$. Persamaan (2.9) dapat ditulis dalam bentuk matriks menjadi (Hardle, 1990)

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{T}) + \boldsymbol{\varepsilon} \tag{2.10}$$

Menurut Eubank (1988), fungsi regresi f dapat diestimasi menggunakan regresi nonparametrik. Pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi, karena data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti (Eubank, 1988).

2.5. Regresi Semiparametrik

Model regresi semiparametrik merupakan gabungan antara regresi parametrik dan regresi nonparametrik. Diberikan bentuk model regresi semiparametrik sebagai berikut :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + f(t_i) + \varepsilon_i,\ i = 1,2,\ldots,n \tag{2.11}$$

Persamaan (2.11) dapat ditulis menjadi

$$\boldsymbol{y} = \mathbf{X\beta} + \boldsymbol{f}(\boldsymbol{T}) + \boldsymbol{\varepsilon} \tag{2.12}$$

dimana $\boldsymbol{y}$ adalah vektor variabel respon yang berukuran $nx1$, $\mathbf{X}$ adalah matriks variabel prediktor untuk komponen parametrik berukuran $n\,x\,(p+1)$, $\boldsymbol{T}$ memuat variabel prediktor komponen nonparametrik, $\boldsymbol{\beta}$ adalah vektor yang berisi parameter regresi berukuran $(p+1)x\,1$, $\boldsymbol{f}$ adalah vektor dari fungsi regresi yang bentuk kurvanya tidak diketahui, $\boldsymbol{\varepsilon}$ adalah vektor error random yang berdistribusi normal dengan mean nol dan variansi $\sigma^2\boldsymbol{I}$ (Salam, 2013).

Matriks pada persamaan (2.12) dimana komponen parametriknya mempunyai hubungan linier dengan variabel respon dan terdapat komponen nonparametrik juga disebut sebagai *Partial Linear Model* (Model Linier Parsial) (Hardle *et al*., 2000). Estimasi model regresi semiparametrik dapat diselesaikan dengan menggunakan metode WLS (*Weighted Least Square*) (Salam, 2013).

2.6. Deret Fourier

Deret Fourier adalah fungsi polinomial trigonometri yang mempunyai tingkat fleksibilitas untuk mengatasi data yang mempunyai pola berulang. Diberikan kurva regresi yang termuat di dalam ruang $C(0,\pi)$ yaitu $f \in C(0,\pi)$, maka estimasi untuk f dapat diselesaikan dengan mengggunakan *Penalized Least Square* atau Kuadrat Terkecil Terpenalti dengan meminimumkan

$$\min_{f \in C(0,\pi)} \left[n^{-1} \sum_{i=1}^{n} (y - f(t))^2 + \lambda \int_0^{\pi} \frac{2}{\pi} (f''(t))^2 dt \right] \tag{2.13}$$

dengan λ merupakan parameter penghalus, dan

$$f(t) = \gamma t + \frac{\alpha_0}{2} + \sum_{k=1}^{K} (\alpha_k \cos kt) \tag{2.14}$$

dimana $\gamma, \alpha_0, \alpha_k, k = 1,2,\ldots,K$ adalah parameter – parameter model. Bilodeau (1992) memberikan estimator untuk kurva regresi nonparametrik Deret Fourier satu respon berbentuk :

$$\hat{F}_{\omega}(t) = \hat{\gamma}(\omega)t + \frac{\hat{\alpha}_0(\omega)}{2} + \sum_{k=1}^{K} (\hat{\alpha}_k(\omega) \cos kt) \tag{2.15}$$

dengan $\widehat{\boldsymbol{\alpha}}(\omega) = (\hat{\gamma}(\omega)\,, \hat{\alpha}_0\,(\omega), \hat{\alpha}_1\,(\omega), \ldots, \hat{\alpha}_k\,(\omega))'$ diperoleh dari persamaan

$$\widehat{\boldsymbol{\alpha}}(\omega) = (n^{-1}\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{\omega G})^{-1} n^{-1} \boldsymbol{X}'\boldsymbol{y}.$$

$\boldsymbol{G}$ merupakan matriks diagonal dan $\boldsymbol{X}$ suatu matriks koefisien (Bilodeau, 1992). Tripena dan Budiantara (2006), menyatakan bahwa apabila *error* random dari model regresi nonparametrik satu respon berdistribusi normal, maka estimator Deret Fourier yang diberikan oleh Bilodeau (1992), mempuyai sifat (i). merupakan estimator bias untuk kurva regresi nonparametrik $f(t)$. (ii).merupakan kelas estimator linier dalam observasi, dan (iii). masing-masing berdistribusi normal (Tripena dan Budiantara, 2006). Dalam beberapa persoalan, sering dijumpai model regresi yang memiliki respon lebih dari satu, misalnya birespon, dimana pola kurva regresinya dapat didekati dengan bentuk trigonomeri. Untuk kasus tersebut secara teoritis estimator Deret Fourier dalam regresi nonparametrik birespon merupakan generalisasi dari estimator Deret Fourier dalam regresi nonparametrik satu respon.

2.7. Regresi Semiparametrik Birespon

Regresi semiparametrik birespon adalah model regresi yang bertujuan untuk mengetahui hubungan antara dua buah variabel respon dengan variabel prediktor dimana antara variabel respon dan prediktor ada yang polanya diketahui, dan ada yang tidak diketahui polanya. Dalam hal ini antar variabel respon saling berkorelasi. Diberikan model regresi semiparametrik birespon sebagai berikut :

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \cdots + \beta_{p1}x_{pi} + f_1(t_i) + \varepsilon_{1i}$$
$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \cdots + \beta_{p2}x_{pi} + f_2(t_i) + \varepsilon_{2i} \;\; ; \;\; i = 1,2,\dots,n \qquad (2.16)$$

Model persamaan (2.16) ini dapat juga dinyatakan dalam bentuk matriks dengan substitusi persamaan (2.14) untuk komponen nonparametriknya.

2.8. Pengujian Dependensi Variabel Respon

Dalam regresi yang melibatkan respon lebih dari satu, antar variabel respon saling dependen. Untuk menguji dependensi dari q variabel respon, $q = 2,3$ dan seterusnya dapat dilakukan uji *Bartlett Sphericity*. Uji *Bartlett Sphericity* sering digunakan dalam analisis multivariat seperti regresi multivariat, analisis faktor, analisis komponen utama, dan sebagainya. Pengujian ini memiliki kelebihan yaitu dapat dilakukan secara serentak karena pengujian ini melibatkan matriks korelasi yang berukuran $q \times q$, dan sesuai untuk sampel besar. Pengujian ini menunjukkan seberapa jauh $|\mathbf{R}| \neq 1$, Jika selisih antara 1 dengan determinan matriks korelasi semakin besar maka dapat menggambarkan antar variabel dependen. Pengujian ini diperoleh dengan menggunakan *Likelihood Ratio* (Morrison, 2005) sebagai berikut :

Hipotesis :

H_0 : **R**=**I** (antar variabel respon independen)

H_1 : **R**≠**I** (antar variabel respon dependen)

Statistik Uji :

$$\chi^2_{hitung} = -\left\{n - 1 - \frac{2q+5}{6}\right\} ln|\mathbf{R}|$$

dimana q adalah jumlah variabel respon dan $|\mathbf{R}|$ adalah nilai determinan matriks korelasi dari masing–masing variabel respon berukuran $q \times q$. **I** adalah matriks identitas orde q. Struktur matriks **R** diberikan sebagai berikut**:**

$$\mathbf{R} = \begin{bmatrix} r_{y1,y1} & r_{y1,y2} & \cdots & r_{y1,yq} \\ r_{y2,y1} & r_{y2,y2} & \cdots & r_{y2,yq} \\ \vdots & \vdots & \cdots & \vdots \\ r_{yq,y1} & r_{yq,y2} & \cdots & r_{yq,yq} \end{bmatrix}$$

dengan

$r_{yi,yj}$=1 untuk $i = j$.

$$r_{yi,yj} = \frac{\sum_{l=1}^{n}(y_{il} - \overline{y}_i)\left(y_{jl} - \overline{y}_j\right)}{\sqrt{\sum_{l=1}^{n}(y_{il} - \overline{y}_i)^2 \sum_{l=1}^{n}\left(y_{jl} - \overline{y}_j\right)^2}}, i \neq j, i = 1,2,\dots,q, j = 1,2,\dots,q$$

$\chi^2_{tabel} = \chi^2_{\alpha,v}$ dimana α merupakan tingkat signifikansi, dan $v = \frac{1}{2}q(q-1)$ adalah derajat kebebasan.

Kriteria Uji :

- Jika $\chi^2_{hitung} \leq \chi^2_{tabel}$ maka H_0 gagal ditolak sehingga antar variabel respon independen.
- Jika $\chi^2_{hitung} > \chi^2_{tabel}$ maka H_0 ditolak sehingga antar variabel respon dependen.

Berdasarkan kriteria uji tersebut diharapkan untuk menolak H_0 sehingga antar variabel respon tersebut saling dependen (Morrison, 2005).

2.9. Pemilihan Parameter Osilasi Optimal dengan GCV

Pada pemodelan regresi nonparametrik dengan menggunakan Deret Fourier, hal yang perlu diperhatikan adalah menentukan nilai K atau parameter osilasi. Pemilihan nilai parameter osilasi menurut Bidermann, Datte, dan Hoffmann (2009) harus dilakukan seoptimal mungkin (Bidermann *et all.*, 2009). Penentuan parameter osilasi optimal bisa menggunakan metode GCV (*Generalized Cross Validation*). Penentuan K optimal akan menghasilkan nilai R^2 yang tinggi (Wahba, 1990). Untuk memperoleh parameter osilasi optimal dapat dilihat dari nilai GCV yang paling minimum. Metode GCV secara umum didefinisikan sebagai berikut (Eubank, 1988).

$$GCV\ (K_1, \dots, K_r) = \frac{MSE\ (K_1, \dots, K_r)}{(n^{-1} trace(\boldsymbol{I} - \boldsymbol{C}(K_1, \dots, K_r)))^2} \tag{2.17}$$

dimana

$$MSE\ (K_1, \dots, K_r) = n^{-1} \boldsymbol{Y}' \left(\boldsymbol{I} - \boldsymbol{C}(K_1, \dots, K_r)\right)' \left(\boldsymbol{I} - \boldsymbol{C}(K_1, \dots, K_r)\right)\boldsymbol{Y} \tag{2.18}$$

$\boldsymbol{C}(K_1, \dots, K_r)$ adalah matriks heat yang memuat parameter osilasi, r adalah banyaknya variabel komponen nonparametrik (Tripena, 2013).

2.10. Koefisien Determinasi dan MSE

Jika nilai koefisien determinasi atau R^2 semakin besar, maka kontribusi prediktor terhadap variabel respon semakin besar. Menurut Eubank (1988), rumus koefisien determinasi diberikan sebagai berikut :

$$R^2 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n_1} (\hat{y}_{ij} - \bar{y})^2}{\sum_{i=1}^{n} \sum_{j=1}^{n_1} (y_{ij} - \bar{y})^2} \ ;\ 0 \leq R^2 \leq 1 \tag{2.19}$$

dimana y_{ij} adalah respon ke $- \ i$ amatan ke $- \ j$, $\hat{y}_{ij}$ adalah estimasi respon ke $- \ i$ amatan ke $- \ j$ dan $\bar{y}$ adalah rata – rata respon. *Mean Square Error* (MSE) adalah nilai dugaan dari variansi residual. Model terbaik adalah model dengan nilai MSE minimum.

2.11. Uji Linieritas

Uji linieritas diperlukan untuk mengetahui pola hubungan tiap variabel prediktor dengan variabel responnya.

Hipotesis :

H_0 : Terdapat hubungan linier antara prediktor dan respon.

H_1 : Tidak terdapat hubungan linier antara prediktor dan respon.

Statistik uji yang digunakan adalah nilai p-value dari tabel Anova antara prediktor dan respon dengan tingkat signifikansi α = 5%.

- Jika *p-value* > α maka H_0 gagal ditolak sehingga terdapat hubungan linier antara prediktor dan respon.
- Jika *p-value* < α maka H_0 ditolak sehingga tidak terdapat hubungan linier antara prediktor dan respon. (Petrovic, 2012)

2.12. Metode *Weighted Least Square* (WLS)

Metode *least square* merupakan prosedur yang dilakukan untuk estimasi parameter, menentukan model yang paling sesuai untuk suatu himpunan data, serta mengkarakterisasi sifat statistik dari estimasi yang diperoleh. Metode *least square* yang sederhana dalam pengunaannya adalah metode OLS (*Ordinary Least Square*), namun dalam kasus pemodelan regresi dengan respon lebih dari satu, metode ini tidak bisa digunakan karena metode OLS tidak mengandung unsur korelasi dan kovariansi. Untuk itu diperlukan metode WLS (*Weighted Least Square*) dimana terdapat matriks pembobot **W**.

$$\mathbf{W} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 & \vdots & \sigma_{12} & \sigma_{12} & \dots & \sigma_{12} \\ 0 & \sigma_1^2 & \dots & \vdots & \vdots & \sigma_{12} & \sigma_{12} & \dots & \sigma_{12} \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sigma_1^2 & \vdots & \sigma_{12} & \sigma_{12} & \dots & \sigma_{12} \\ \dots & \dots & \dots & \dots & \vdots & \dots & \dots & \dots & \dots \\ \sigma_{21} & \sigma_{21} & \dots & \sigma_{21} & \vdots & \sigma_2^2 & 0 & \dots & 0 \\ \sigma_{21} & \sigma_{21} & \dots & \sigma_{21} & \vdots & 0 & \sigma_2^2 & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ \sigma_{21} & \sigma_{21} & \dots & \sigma_{21} & \vdots & 0 & 0 & \dots & \sigma_2^2 \end{pmatrix}$$

Dengan menggunakan prosedur yang sama seperti OLS, metode ini mendefinisikan estimasi parameter sebagai suatu nilai yang meminimumkan jumlah kuadrat antara pengamatan dan model yang disebut jumlah kuadrat error yang didefinisikan sebagai (Searle, 1971) :

$$\boldsymbol{\varepsilon}^{\mathbf{T}}\mathbf{W}\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^{\mathbf{T}}\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{2.20}$$

2.13. Beberapa Teorema Penunjang dalam Aljabar Matriks

Berikut ini diberikan beberapa pengertian dasar yang digunakan dalam proses mendapatkan estimator model regresi semiparametrik birespon dengan pendekatan Deret Fourier :

Teorema 2.13.1

Diberikan matriks **A** dan **B** yang saling *close form*, maka berlaku sifat – sifat sebagai berikut :

a. Jika matriks **A** simetris, maka $\mathbf{A}^{\mathbf{T}}=\mathbf{A}$.
b. $(\mathbf{A}+\mathbf{B})^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}} + \mathbf{B}^{\mathbf{T}}$.
c. $(\mathbf{AB})^{\mathbf{T}} = \mathbf{B}^{\mathbf{T}}\,\mathbf{A}^{\mathbf{T}}$ ■

Teorema berikutnya berkaitan dengan diferensial pada matriks dan vektor yang diperlukan.

Teorema 2.13.2

Diberikan vektor **a** dan **x**, dimana $\mathbf{a}^{\mathbf{T}}\mathbf{x} = \mathbf{x}^{\mathbf{T}}\mathbf{a}$, $\mathbf{a}^{\mathbf{T}} = (a_1, a_2, \dots, a_p)$ memuat konstanta, dengan ini berlaku

$$\frac{\partial(\mathbf{a}^{\mathbf{T}}\mathbf{x})}{\partial\mathbf{x}} = \frac{\partial(\mathbf{x}^{\mathbf{T}}\mathbf{a})}{\partial\mathbf{x}} = \mathbf{a}$$ ■

Teorema 2.13.3

Diberikan vektor **x** dan **A** merupakan suatu matriks simetri, maka

$$\frac{\partial(\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x})}{\partial\mathbf{x}} = 2\mathbf{A}\mathbf{x}$$ ■

Teorema – teorema tersebut diambil dari (Rencher dan Schaalje, 2008).

2.14. Indeks Komposit Pembangunan Sosial - Ekonomi

Perdebatan tentang indikator komposit pembangunan sosial-ekonomi sudah sejak lama terjadi, sebab indikator ini tidak hanya bertolak ukur pada pendapatan per kapita. Pendapatan perkapita sebagai indikator pembangunan tidak disetujui oleh kalangan ekonomi maupun non-ekonomi yang melihat ketidakakuratan indikator tersebut, sehingga muncul beberapa indikator baru. Morris (1979) membangun *the Physical Quality of Life Index* (*PQLI*), sedangkan United Nation Development Program (UNDP) membangun *Human Development Index* (*HDI*) atau Indeks Pembangunan Manusia (IPM) yang kini banyak digunakan oleh negara di dunia dengan landasan yang dibangun oleh Haq (1996).

Indeks Pembangunan Manusia (IPM) merupakan indikator komposit yang mengukur tiga dimensi pokok pembangunan manusia dan dinilai mampu mencerminkan kemampuan dasar (*basic capabilities*) penduduk. Ketiga kemampuan dasar itu adalah panjang umur dan sehat, berpengetahuan dan berketerampilan, serta akses terhadap sumber daya yang dibutuhkan untuk mencapai standar hidup layak. UNDP mendefinisikan pembangunan manusia sebagai suatu proses untuk memperluas pilihan-pilihan bagi penduduk dalam hal pendapatan, kesehatan, pendidikan, lingkungan fisik, dan sebagainya. Empat hal pokok yang perlu di perhatikan dalam pembangunan manusia adalah produktivitas, pemerataan, kesinambungan, pemberdayaan (UNDP, 1995). Titik berat pembangunan nasional Indonesia sesungguhnya sudah menganut konsep tersebut. IPM dihitung dengan rumus sebagai berikut:

$$IPM = (IL + IK + IDL) / 3 \qquad (2.21)$$

dimana

- IL : Indeks yang berasal dari angka harapan hidup (AHH).
- IK : Indeks yang berasal MYS dan AMH, dimana IK = 2/3 AMH + 1/3 MYS. AMH adalah Angka Melek Huruf, MYS adalah rata – rata lama sekolah.
- IDL : Indeks yang berasal dari GNP atau pengeluaran perkapita tiap bulan.

Perhitungan tiap komponen indeks penyusun IPM adalah

$$IX = (X - X_{min})/(X_{max} - X_{min}) \quad (2.22)$$

dengan

IX = Indeks komponen X dari IPM

X = komponen X dari IPM

X_{max} = nilai maksimum komponen X

X_{min} = nilai minimum komponen X

nilai komponen maksimum komponen X dan minimumnya diberikan dalam Tabel 2.1 (UNDP, 1995). Tabel 2.1 berikut adalah nilai maksimum dan minimum komponen IPM yang ditetapkan oleh UNDP.

Tabel 2.1 Nilai Maksimum dan Minimum Komponen IPM

Indikator	Minimum	Maksimum
AHH	25	85
AMH	0	100
LS	0	15
GDP	300.000 (1996), 360.000 (1999)	732,72

Indikator kesejahteraan lain adalah Indeks Mutu Hidup (IMH) atau PQLI (*Physical Quality of Life Index*) yang dikenalkan oleh Morris. IMH adalah indeks pembangunan kombinasi dari tiga indikator yaitu angka kematian bayi, angka harapan hidup dari indikator kesehatan, dan angka melek huruf dari indikator pendidikan. IMH sesuai untuk mengukur capaian pembangunan manusia suatu daerah yang pendapatan perkapitanya rendah. IMH digunakan untuk mengukur kualitas hidup atau kesejahteraan suatu wilayah. David Morris (1970) mengembangkan IMH untuk *Overseas Development Council*, sebagai salah satu solusi atas ketidakpuasan ketika menggunakan GNP sebagai indikator pembangunan (Roger Nosal dan Tom Schultz, 2008). Rumus Morris (Morris, 1979) yang digunakan untuk mengukur IMH adalah sebagai berikut :

$$IMH = \frac{AMH + (166 - AKB)0{,}625 + (AHH - 42)2{,}7}{3} \quad (2.23)$$

“halaman ini sengaja dikosongkan”

BAB 3
METODE PENELITIAN

3.1 Sumber Data dan Struktur Data

Data yang digunakan dalam peneltian ini merupakan data sekunder tentang IPM (Indeks Pembangunan Manusia), Indeks Kesehatan Masyarakat, Indeks Pendidikan, PDRB (Prooduk Domestik Regional Bruto), Indeks PPP (*Purchasing Power Parity*), TPAK (Tingkat Partisipasi Angkatan Kerja), dan Persentase Kemiskinan di Jawa Timur tahun 2012. Untuk menghitung IMH, diperlukan data AMH (Angka Melek Huruf), AHH (Angka Harapan Hidup), dan AKB (Angka Kematian Bayi) di Jawa Timur tahun 2012. Data diambil dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Struktur data penelitian ini diberikan sebagai berikut :

Tabel 3.1. Struktur Data Penelitian

i	Daerah	Y_1	Y_2	X_1	X_2	...	X_p	t_1	t_2	...	t_r
1	Pacitan	$y_{1,1}$	$y_{2,1}$	$x_{1,1}$	$x_{2,1}$	...	$x_{p,1}$	$t_{1,1}$	$t_{2,1}$	...	$t_{r,1}$
2	Ponorogo	$y_{1,2}$	$y_{2,2}$	$x_{1,2}$	$x_{2,2}$	...	$x_{p,2}$	$t_{1,2}$	$t_{2,2}$	...	$t_{r,2}$
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	...	⋮
n=38	Kota Batu	$y_{1,38}$	$y_{2,38}$	$x_{1,38}$	$x_{2,38}$	...	$x_{p,38}$	$t_{1,38}$	$t_{2,38}$	...	$t_{r,38}$
		Respon		Komponen Parametrik				Komponen Nonparametrik			

3.2 Variabel Penelitian

Variabel dalam penelitian ini terdiri atas dua variabel respon dan enam variabel prediktor dengan tipe data kontinu sebagai berikut beserta definisi operasional dari Badan Pusat Statistik (BPS) :

- Y_1 adalah variabel respon nilai IPM wilayah di Jawa Timur. Indeks pembangunan manusia mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup.

- Y_2 adalah variabel respon nilai IMH wilayah di Jawa Timur. Indeks mutu hidup mengukur capaian pembangunan manusia berbasis komponen pendidikan dan kesehatan.
- Z_1 adalah variabel prediktor Indeks Kesehatan Masyarakat wilayah di Jawa Timur. Indeks kesehatan masyarakat adalah indeks yang diukur berdasarkan indikator – indikator kesehatan.
- Z_2 adalah variabel prediktor Indeks Pendidikan wilayah di Jawa Timur. Indeks pendidikan adalah indeks yang diukur berdasarkan indikator – indikator pendidikan.
- Z_3 adalah variabel prediktor Produk Domestik Regional Bruto (PDRB). Produk domestik regional bruto adalah jumlah nilai tambah yang dihasilkan untuk seluruh wilayah usaha dalam suatu wilayah atau merupakan jumlah seluruh nilai barang dan jasa akhir yang dihasilkan seluruh unit ekonomi di suatu wilayah.
- Z_4 adalah variabel prediktor Indeks PPP (*Purchasing Power Parity*) wilayah di Jawa Timur. Indeks PPP adalah indeks paritas daya beli yang diukur berdasarkan daya beli masyarakat terhadap kebutuhan pokok.
- Z_5 adalah variabel prediktor Tingkat Partisipasi Angkatan Kerja (TPAK). TPAK adalah persentase angkatan kerja terhadap jumlah penduduk usia kerja.
- Z_6 adalah variabel prediktor Persentase Kemiskinan. Persentase kemiskinan adalah persentase penduduk yang berada dibawah garis kemiskinan.

(Badan Pusat Statistik, 2013)

Berdasarkan *scatter plot* dan uji linieritas antara variabel respon dan prediktor data berpasangan (z_i, y_i), jika data berpola atau linier menurut uji linieritas, maka prediktor digolongkan dalam variabel prediktor parametrik yang kemudian dilambangkan dengan X_i. Jika data tidak berpola atau tidak linier menurut uji linieritas, maka prediktor digolongkan dalam variabel prediktor nonparametrik yang kemudian dilambangkan dengan t_i.

3.3 Langkah-langkah Penelitian

Berikut adalah langkah – langkah penelitian yang akan dilakukan:

1. Mendapatkan estimasi model regresi semiparametrik Deret Fourier birespon dengan langkah-langkah sebagai berikut :

a. Membuat model regresi semiparametrik birespon. Diberikan variabel respon Y_1 dan Y_2 dengan variabel komponen parametrik $X_1, X_2 \ldots, X_p$, dan variabel komponen nonparametrik $t_1, t_2 \ldots, t_R$:

$$y_{ji} = \beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi} + \sum_{r=1}^{R} f_j(t_{ri}) + \varepsilon_{ji}$$

dimana $i = 1,2,\ldots,n$,dan $j = 1,2$.

b. Mendekati komponen nonparametrik dengan fungsi Deret Fourier

$$f_j(t_{ri}) = \gamma_{jr} t_{ri} + \frac{\alpha_{0jr}}{2} + \sum_{k=1}^{K} (\alpha_{kjr} \cos k t_{ri})$$

dimana $i = 1,2,\ldots,n$, dan j=1,2.

c. Model regresi semiparametrik birespon ditulis dalam bentuk matriks

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{t}) + \boldsymbol{\varepsilon}$$

dimana $\boldsymbol{y} = \left[\boldsymbol{y}_1^T \vdots \boldsymbol{y}_2^T\right]^T$, dan $\boldsymbol{\varepsilon} = \left[\boldsymbol{\varepsilon}_1^T \vdots \boldsymbol{\varepsilon}_2^T\right]^T$ adalah vektor yang berukuran $2n \times 1$.

d. Membentuk *goodness of fit* :

$$R(\boldsymbol{f}) = \left(\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{t})\right)^T \boldsymbol{W} \left(\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{t})\right)$$

e. Mendapatkan estimasi model dengan menyelesaikan optimasi :

$$\min_{f \in C(0,\pi)} [R(\boldsymbol{f})]$$

Optimasi diselesaikan menggunakan derivatif parsial.

2. Memodelkan hubungan antara IPM dan IMH di Jawa Timur dengan variabel – variabel yang mempengaruhinya dengan langkah-langkah sebagai berikut :

a. Melakukan analisis deskriptif untuk mengetahui gambaran umum tentang indikator pembangunan dan kesejahteraan di Jawa Timur beserta komponen penentunya.

b. Menguji korelasi antara respon pertama dan respon kedua.

c. Melakukan analisis regresi parametrik birespon untuk mengetahui kekurangan metode tersebut terhadap data yang diteliti, sehingga menjadi dasar untuk menggunakan regresi semiparametrik birespon dengan pendekatan Deret Fourier.

d. Membuat *scatter plot* dan melakukan uji linieritas untuk mengidentifikasi pola data mana yang tergolong komponen parametrik dan komponen nonparametrik menurut interpretasi *scatter plot* dan uji linieritas.
e. Membuat program penentuan parameter osilasi optimal berdasarkan kriteria GCV untuk komponen nonparametrik dimana

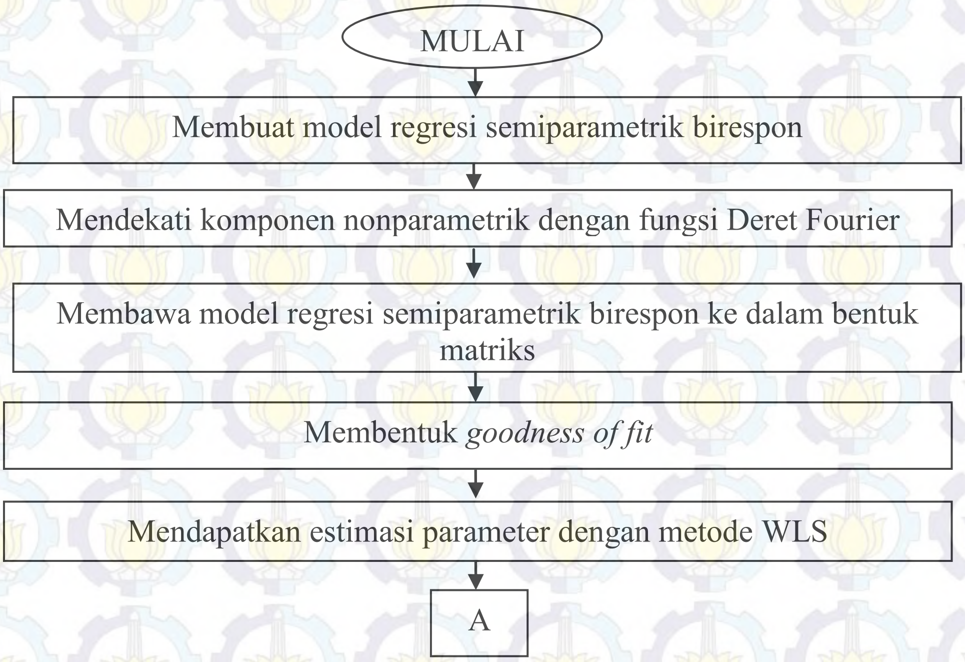
$$GCV\ (K_1, K_2, \dots, K_r) = \frac{MSE\ (K_1, K_2, \dots, K_r)}{((2n)^{-1} trace(\boldsymbol{I} - \boldsymbol{C}(K_1, K_2, \dots, K_r)))^2}$$

dengan

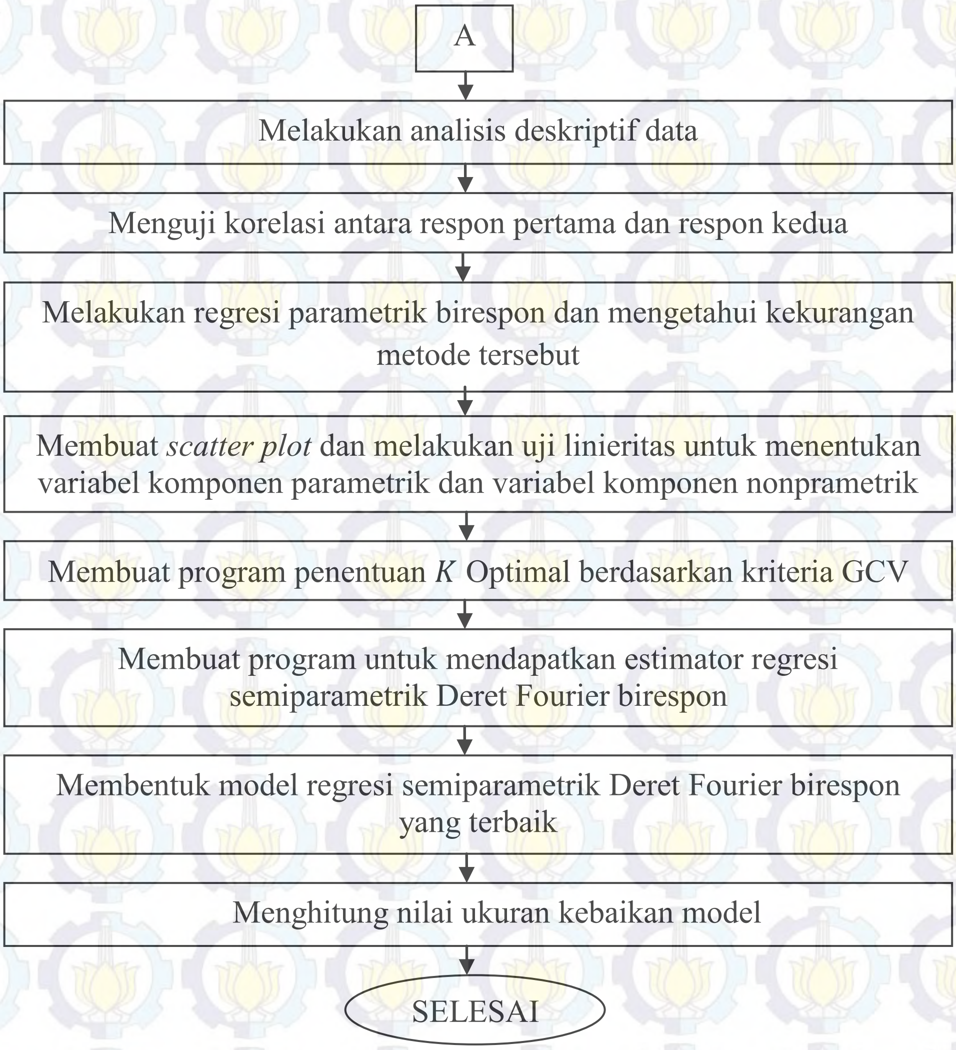
$$MSE\ (K_1, \dots, K_r) = (2n)^{-1} \boldsymbol{Y}' \left(\boldsymbol{I} - \boldsymbol{C}(K_1, K_2, \dots, K_r)\right)' \left(\boldsymbol{I} - \boldsymbol{C}(K_1, K_2, \dots, K_r)\right)\boldsymbol{Y}$$

f. Membuat program untuk mendapatkan estimator regresi semiparametrik Deret Fourier birespon.
g. Membentuk model regresi semiparametrik Deret Fourier birespon yang terbaik.
h. Menghitung nilai ukuran kebaikan model .

Adapun diagram tahap penelitian disajikan sebagai berikut



Gambar 3.1 Diagram Alir Penelitian Tujuan Pertama



Gambar 3.2 Diagram Alir Penelitian Tujuan Kedua

"halaman ini sengaja dikosongkan"

BAB 4
HASIL DAN PEMBAHASAN

Pada Bab 4 ini dibahas mengenai langkah-langkah untuk mendapatkan estimasi parameter dari model regresi semiparametrik birespon dengan pendekatan Deret Fourier. Estimasi parameter tersebut diperoleh dengan menggunakan metode optimasi WLS (*Weighted Least Square*). Sementara untuk pemilihan parameter osilasi yang optimal menggunakan metode GCV (*Generalized Cross Validation*). Setelah didapatkan estimasi parameter maka selanjutnya adalah menerapkan model regresi semiparametrik birespon menggunakan pendekatan Deret Fourier ke dalam permasalahan indikator kesejahteraan daerah di Jawa Timur dalam hal ini IPM dan IMH.

4.1 Estimasi Model Regresi Semiparametrik Birespon dengan Pendekatan Deret Fourier

Diberikan data berpasangan $(x_1, x_2 \ldots, x_p, t_1, t_2 \ldots, t_R,\ y_1, y_2)$. Variabel prediktor $x_1, x_2 \ldots, x_p$ mengikuti pola tertentu sedangkan variabel prediktor $t_1, t_2 \ldots, t_R$ tidak mengikuti pola tertentu. Model regresi semiparametrik birepon yang memuat variabel tersebut dinyatakan sebagai berikut.

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \beta_{21}x_{2i} + \ldots + \beta_{p1}x_{pi} + \sum_{r=1}^{R} f_1(t_{ri}) + \varepsilon_{1i}$$

$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \beta_{22}x_{2i} + \ldots + \beta_{p2}x_{pi} + \sum_{r=1}^{R} f_2(t_{ri}) + \varepsilon_{2i} \tag{4.1}$$

dimana $i = 1,2,\ldots,n$, dan $p = 1,2,\ldots,P$.

Persamaan di atas dapat ditulis dalam bentuk matriks masing–masing berukuran $2n \times 1$ berikut

$$\begin{bmatrix} \mathbf{y_1} \\ \cdots \\ \mathbf{y_2} \end{bmatrix} = \begin{bmatrix} \mathbf{X_1\beta_1} \\ \cdots \\ \mathbf{X_2\beta_2} \end{bmatrix} + \begin{bmatrix} \mathbf{f_1} \\ \cdots \\ \mathbf{f_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \cdots \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} \tag{4.2}$$

dengan

$$\mathbf{y_1} = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \end{bmatrix}, \ \mathbf{y}_2 = \begin{bmatrix} y_{21} \\ y_{22} \\ \vdots \\ y_{2n} \end{bmatrix}$$ memuat variabel respon

$$\boldsymbol{\beta_1} = \begin{bmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{p1} \end{bmatrix}, \ \boldsymbol{\beta_2} = \begin{bmatrix} \beta_{02} \\ \beta_{12} \\ \beta_{22} \\ \vdots \\ \beta_{p2} \end{bmatrix}$$ memuat parameter komponen parametrik

$$\mathbf{f_1} = \begin{bmatrix} f_1(t_{11}) + f_1(t_{21}) + \ldots + f_1(t_{r1}) \\ f_1(t_{12}) + f_1(t_{22}) + \ldots + f_1(t_{r2}) \\ \vdots \\ f_1(t_{1n}) + f_1(t_{2n}) + \ldots + f_1(t_{rn}) \end{bmatrix}, \ \mathbf{f_2} = \begin{bmatrix} f_2(t_{11}) + f_2(t_{21}) + \ldots + f_2(t_{r1}) \\ f_2(t_{12}) + f_2(t_{22}) + \ldots + f_2(t_{r2}) \\ \vdots \\ f_2(t_{1n}) + f_2(t_{2n}) + \ldots + f_2(t_{rn}) \end{bmatrix}$$

yang memuat komponen nonparametrik, dan vektor error random

$$\boldsymbol{\varepsilon_1} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n} \end{bmatrix}, \ \boldsymbol{\varepsilon_2} = \begin{bmatrix} \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2n} \end{bmatrix}$$

Vektor–vektor tersebut berukuran $n \times 1$. Selanjutnya diberikan matriks yang memuat variabel prediktor komponen parametrik.

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_{111} & x_{211} & \cdots & x_{p11} \\ 1 & x_{121} & x_{221} & \cdots & x_{p21} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n1} & x_{2n1} & \cdots & x_{pn1} \end{bmatrix}, \ \mathbf{X}_2 = \begin{bmatrix} 1 & x_{112} & x_{212} & \cdots & x_{p12} \\ 1 & x_{121} & x_{222} & \cdots & x_{p22} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n2} & x_{2n2} & \cdots & x_{pn2} \end{bmatrix}$$

yang masing – masing berukuran $n \times (p+1)$.

Persamaan matriks (4.2) secara umum dapat disajikan dalam bentuk

$$\mathbf{y}_{2n\times1} = \mathbf{X}_{2n\times2(p+1)}\boldsymbol{\beta}_{2(p+1)\times1} + \mathbf{f}_{2n\times1} + \boldsymbol{\varepsilon}_{2n\times1} \tag{4.3}$$

dimana $\mathbf{f}$ memuat fungsi–fungsi yang diasumsikan bentuk pola kurva regresinya tidak diketahui sehingga didekati dengan fungsi Deret Fourier sebagai berikut

$$f_1(t_{ri}) = \gamma_{1r} t_{ri} + \frac{\alpha_{01r}}{2} + \sum_{k=1}^{K} (\alpha_{k1r} \cos k t_{ri})$$

$$f_2(t_{ri}) = \gamma_{2r} t_{ri} + \frac{\alpha_{02r}}{2} + \sum_{k=1}^{K} (\alpha_{k2r} \cos k t_{ri}); i = 1,2,...,n, r = 1,2,...,R. \quad (4.4)$$

dengan K adalah parameter osilasi. **X** adalah suatu matriks dengan komponen

$$\mathbf{X} = \begin{bmatrix} \mathbf{X_1} & \vdots & \mathbf{O} \\ \cdots & \vdots & \cdots \\ \mathbf{O} & \vdots & \mathbf{X_2} \end{bmatrix}$$

dimana **O** adalah matriks yang semua elemennya nol.

Selanjutnya untuk mengetahui matriks komponen nonparametrik dari model regresi semiparametrik birespon dengan pendekatan Deret Fourier, persamaan (4.3) diubah menjadi.

$$\mathbf{y}_{2n\times 1} = \mathbf{X}_{2n\times 2(p+1)} \boldsymbol{\beta}_{2(p+1)\times 1} + \mathbf{T}_{2n\times(2(r(k+1)+1))} \boldsymbol{\eta}_{(2(r(k+1)+1))\times 1} + \boldsymbol{\varepsilon}_{2n\times 1}$$

dimana

$$\mathbf{T} = \begin{bmatrix} \mathbf{T_1} & \vdots & \mathbf{O} \\ \cdots & \vdots & \cdots \\ \mathbf{O} & \vdots & \mathbf{T_2} \end{bmatrix}$$

dengan

$$\mathbf{T}_1 = \begin{bmatrix} 1 & t_{111} & \cos t_{111} & \cdots & \cos K t_{111} & t_{211} & \cos t_{211} & \cdots & \cos K t_{211} & t_{r11} & \cos t_{r11} & \cdots & \cos K t_{r11} \\ 1 & t_{121} & \cos t_{121} & \cdots & \cos K t_{121} & t_{221} & \cos t_{221} & \cdots & \cos K t_{221} & t_{r21} & \cos t_{r21} & \cdots & \cos K t_{r21} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{1n1} & \cos t_{1n1} & \cdots & \cos K t_{1n1} & t_{2n1} & \cos t_{2n1} & \cdots & \cos K t_{2n1} & t_{rn1} & \cos t_{rn1} & \cdots & \cos K t_{rn1} \end{bmatrix}$$

$$\mathbf{T}_2 = \begin{bmatrix} 1 & t_{112} & \cos t_{112} & \cdots & \cos K t_{112} & t_{212} & \cos t_{212} & \cdots & \cos K t_{212} & t_{r12} & \cos t_{r12} & \cdots & \cos K t_{r12} \\ 1 & t_{122} & \cos t_{122} & \cdots & \cos K t_{122} & t_{222} & \cos t_{222} & \cdots & \cos K t_{222} & t_{r22} & \cos t_{r22} & \cdots & \cos K t_{r22} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{1n2} & \cos t_{1n2} & \cdots & \cos K t_{1n2} & t_{2n2} & \cos t_{2n2} & \cdots & \cos K t_{2n2} & t_{rn2} & \cos t_{rn2} & \cdots & \cos K t_{rn2} \end{bmatrix}$$

adalah matriks yang masing – masing berukuran $n \times [r(k+1)+1]$.

$$\boldsymbol{\eta} = [\boldsymbol{\eta}_1^{\mathbf{T}} \vdots \boldsymbol{\eta}_2^{\mathbf{T}}]^{\mathbf{T}}$$

$$\boldsymbol{\eta}_1 = (\alpha_{01}^*, \gamma_{11}, \alpha_{111}, \alpha_{211}, \ldots, \alpha_{K11}, \gamma_{12}, \alpha_{112}, \alpha_{212}, \ldots, \alpha_{K12}, \ldots, \gamma_{1r}, \alpha_{11r}, \alpha_{21r}, \ldots, \alpha_{K1r})^T$$

$$\boldsymbol{\eta}_2 = (\alpha_{02}^*, \gamma_{21}, \alpha_{121}, \alpha_{221}, \ldots, \alpha_{K21}, \gamma_{22}, \alpha_{122}, \alpha_{222}, \ldots, \alpha_{K22}, \ldots, \gamma_{2r}, \alpha_{12r}, \alpha_{22r}, \ldots, \alpha_{K2r})^T$$

dimana $\alpha_{0j}^* = \frac{r}{2}\alpha_{0j}$

Jika persamaan (4.4) disubstitusikan ke persamaan (4.1) maka model untuk setiap respon ke- j, $j = 1,2$, variabel prediktor komponen parametrik ke- p , variabel prediktor komponen nonparametrik ke-r, dan pengamatan ke-i dapat ditulis sebagai berikut :

$$y_{1i}=\beta_{01}+\beta_{11}x_{1i}+\beta_{21}x_{2i}+\ldots+\beta_{p1}x_{pi}+\sum_{r=1}^{R}(\gamma_{1r}t_{ri}+\frac{\alpha_{01r}}{2}+\sum_{k=1}^{K}(\alpha_{k1r}\cos kt_{ri}))+\varepsilon_{1i}$$

$$y_{2i}=\beta_{02}+\beta_{12}x_{1i}+\beta_{22}x_{2i}+\ldots+\beta_{p2}x_{pi}+\sum_{r=1}^{R}(\gamma_{2r}t_{ri}+\frac{\alpha_{02r}}{2}+\sum_{k=1}^{K}(\alpha_{k2r}\cos kt_{ri}))+\varepsilon_{2i}$$

dimana $i=1,2,\ldots,n$, $p=1,2,\ldots,P$, $r=1,2,\ldots,R$ dan $k=1,2,\ldots,K$. Error random ε_{1i} dan ε_{2i} saling berkorelasi. Error random ε_{1i} dan ε_{2i} memiliki mean nol, variansi masing–masing σ_1^2 dan σ_2^2. Struktur matriks varian kovarian dari error dapat ditulis sebagai berikut:

$$\mathbf{W}=\begin{pmatrix}
\sigma_1^2 & 0 & \ldots & 0 & \vdots & \sigma_{12} & \sigma_{12} & \ldots & \sigma_{12}\\
0 & \sigma_1^2 & \ldots & \vdots & \vdots & \sigma_{12} & \sigma_{12} & \ldots & \sigma_{12}\\
\vdots & \vdots & \ddots & 0 & \vdots & \vdots & \vdots & \ldots & \vdots\\
0 & 0 & \ldots & \sigma_1^2 & \vdots & \sigma_{12} & \sigma_{12} & \ldots & \sigma_{12}\\
\ldots & \ldots & \ldots & \ldots & \vdots & \ldots & \ldots & \ldots & \ldots\\
\sigma_{21} & \sigma_{21} & \ldots & \sigma_{21} & \vdots & \sigma_2^2 & 0 & \ldots & 0\\
\sigma_{21} & \sigma_{21} & \ldots & \sigma_{21} & \vdots & 0 & \sigma_2^2 & \ldots & \vdots\\
\vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ddots & 0\\
\sigma_{21} & \sigma_{21} & \ldots & \sigma_{21} & \vdots & 0 & 0 & \ldots & \sigma_2^2
\end{pmatrix} \tag{4.5}$$

Misalkan $\theta=\sigma_1\sigma_2$, $r=\frac{\sigma_1}{\sigma_2}$ dan korelasi $\rho=\frac{\sigma_{12}}{\sigma_1\sigma_2}$, maka kovariansi σ_{12} dapat ditulis menjadi $\sigma_{12}=\rho\sigma_1\sigma_2=\rho\theta$. Variansi σ_1^2 dan σ_2^2 dapat disajikan menjadi $\sigma_1^2=\sigma_1\sigma_2\left(\frac{\sigma_1}{\sigma_2}\right)=\theta r$, dan $\sigma_2^2=\sigma_1\sigma_2\left(\frac{\sigma_2}{\sigma_1}\right)=\sigma_1\sigma_2\frac{1}{\left(\frac{\sigma_1}{\sigma_2}\right)}=\frac{\theta}{r}$. Dengan demikian matriks **W** menjadi

$$\mathbf{W}=\begin{pmatrix}
\theta r & 0 & \ldots & 0 & \vdots & \theta\rho & \theta\rho & \ldots & \theta\rho\\
0 & \theta r & \ldots & \vdots & \vdots & \theta\rho & \theta\rho & \ldots & \theta\rho\\
\vdots & \vdots & \ddots & 0 & \vdots & \vdots & \vdots & \ldots & \vdots\\
0 & 0 & \ldots & \theta r & \vdots & \theta\rho & \theta\rho & \ldots & \theta\rho\\
\ldots & \ldots & \ldots & \ldots & \vdots & \ldots & \ldots & \ldots & \ldots\\
\theta\rho & \theta\rho & \ldots & \theta\rho & \vdots & \theta/r & 0 & \ldots & 0\\
\theta\rho & \theta\rho & \ldots & \theta\rho & \vdots & 0 & \theta/r & \ldots & \vdots\\
\vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ddots & 0\\
\theta\rho & \theta\rho & \ldots & \theta\rho & \vdots & 0 & 0 & \ldots & \theta/r
\end{pmatrix}$$

$$= \theta \begin{pmatrix} r\boldsymbol{I_n} & \vdots & \rho \boldsymbol{J_{nxn}} \\ \cdots & \vdots & \cdots \\ \rho \boldsymbol{J_{nxn}} & \vdots & \frac{1}{r}\boldsymbol{I_n} \end{pmatrix}$$

dimana $\boldsymbol{I_n}$ adalah matriks identitas orde n dan $\boldsymbol{J_{nxn}}$ adalah matriks yang memuat elemen satu berukuran $n \times n$. Matriks $\mathbf{W}$ digunakan dalam optimasi *Weighted Least Square* (WLS) untuk mendapatkan estimator parameter $\boldsymbol{\beta}$ dan $\boldsymbol{\eta}$. Optimasi WLS dilakukan dengan meminimumkan *goodness of fit* dari model regresi semiparametrik birespon dengan pendekatan Deret Fourier

$$\min_{\boldsymbol{f} \in C(0,\pi)} [\, R(\mathbf{f})] = \min_{\boldsymbol{f} \in C(0,\pi)} \left\{ \left(\mathbf{y} - \mathbf{f}(\mathbf{x},\mathbf{t})\right)^{\mathrm{T}} \mathbf{W}\left(\mathbf{y} - \mathbf{f}(\mathbf{x},\mathbf{t})\right) \right\}$$

$\mathbf{f}$ memuat parameter $\boldsymbol{\beta}$ dan $\boldsymbol{\eta}$ sehingga

$$\min_{\boldsymbol{\beta},\boldsymbol{\eta} \in R^{2(p+1)}} [\, R(\boldsymbol{\beta},\boldsymbol{\eta})] = \min_{\boldsymbol{\beta},\boldsymbol{\eta} \in R^{2(p+1)}} \left\{ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta})^{\mathrm{T}} \mathbf{W}\left(\mathbf{y} - \mathbf{f}(\mathbf{x},\mathbf{t})\right) \right\} \quad (4.6)$$

Dengan melakukan penjabaran pada optimasi persamaan (4.6) dan menerapkan Teorema 2.13.1 , maka didapat

$$\begin{aligned} R(\boldsymbol{\beta},\boldsymbol{\eta}) &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta})^{\mathrm{T}} \mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta}) \\ &= (\mathbf{y}^{\mathrm{T}} - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}} - \boldsymbol{\eta}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}})\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta}) \\ &= (\mathbf{y}^{\mathrm{T}}\mathbf{W} - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{W} - \boldsymbol{\eta}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{W})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta}) \\ &= \mathbf{y}^{\mathrm{T}}\mathbf{W}\mathbf{y} - 2\mathbf{y}^{\mathrm{T}}\mathbf{W}\mathbf{X}\boldsymbol{\beta} - 2\boldsymbol{\eta}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{W}\mathbf{y} + 2\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{T}\boldsymbol{\eta} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{W}\mathbf{T}\boldsymbol{\eta} \end{aligned}$$

Untuk mendapatkan estimator dari parameter $\boldsymbol{\beta}$ dilakukan dengan melakukan derivatif parsial $R(\boldsymbol{\beta},\boldsymbol{\eta})$ terhadap $\boldsymbol{\beta}$. Dalam proses derivatif ini digunakan Teorema 2.13.2 dan Teorema 2.13.3 sebagai konsep dasar.

$$\frac{\partial R(\boldsymbol{\beta},\boldsymbol{\eta})}{\partial \boldsymbol{\beta}} = \mathbf{0}$$

$$\frac{\partial \left\{ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta})^{\mathrm{T}} \mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta}) \right\}}{\partial \boldsymbol{\beta}} = \mathbf{0}$$

$$\begin{aligned} -2\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{y} + 2\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{T}\hat{\boldsymbol{\eta}} + 2\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{0} \\ -\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{y} + \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{T}\hat{\boldsymbol{\eta}} + \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{0} \\ \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{y} - \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{T}\hat{\boldsymbol{\eta}} \end{aligned}$$

diperoleh

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\right)^{-1} \left\{ \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{y} - \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{T}\hat{\boldsymbol{\eta}} \right\} \quad (4.7)$$

Untuk mendapatkan estimator dari parameter $\boldsymbol{\eta}$ dilakukan dengan melakukan derivatif parsial $R(\boldsymbol{\beta},\boldsymbol{\eta})$ terhadap $\boldsymbol{\eta}$. Dalam proses derivatif ini digunakan Teorema 2.13.2 dan Teorema 2.13.3 sebagai konsep dasar.

$$\frac{\partial R(\boldsymbol{\beta},\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \mathbf{0}$$

$$\frac{\partial\{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{T}\boldsymbol{\eta})^{T}\mathbf{W}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{T}\boldsymbol{\eta})\}}{\partial \boldsymbol{\eta}} = \mathbf{0}$$

$$-\mathbf{2T}^{T}\mathbf{Wy}+\mathbf{2T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}+\mathbf{2T}^{T}\mathbf{WT}\hat{\boldsymbol{\eta}} = \mathbf{0}$$

$$-\mathbf{T}^{T}\mathbf{Wy}+\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}+\mathbf{T}^{T}\mathbf{WT}\hat{\boldsymbol{\eta}} = \mathbf{0}$$

$$\mathbf{T}^{T}\mathbf{WT}\hat{\boldsymbol{\eta}} = \mathbf{T}^{T}\mathbf{Wy}-\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}$$

diperoleh

$$\hat{\boldsymbol{\eta}} = \left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\left\{\mathbf{T}^{T}\mathbf{Wy}-\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}\right\} \tag{4.8}$$

Estimator dalam persamaan (4.7) dan (4.8) belum bebas parameter sehingga harus dicari estimator yang bebas dari parameter dengan saling substitusi. Untuk mendapatkan $\hat{\beta}$ yang bebas parameter, substitusi persamaan (4.8) ke persamaan (4.7).

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= \left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\left[\mathbf{X}^{T}\mathbf{Wy}-\mathbf{X}^{T}\mathbf{WT}\left\{\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\left\{\mathbf{T}^{T}\mathbf{Wy}-\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}\right\}\right\}\right] \\
&= \left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\left[\mathbf{X}^{T}\mathbf{Wy}-\mathbf{X}^{T}\mathbf{WT}\left\{\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{Wy}-\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}\right\}\right] \\
&= \left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\left[\mathbf{X}^{T}\mathbf{Wy}-\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{Wy}+\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}\right] \\
&= \left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{Wy}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{Wy}+\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT} \\
&\quad \left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}}
\end{aligned}$$

Selanjutnya suku yang memuat parameter dikelompokkan dalam satu ruas sebagai berikut :

$$\hat{\boldsymbol{\beta}}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{WX}\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{Wy}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{Wy}$$

Persamaan di atas dapat diuraikan menjadi

$$\hat{\boldsymbol{\beta}}(\mathbf{I}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{WX}) = \left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{Wy}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{Wy}$$

Suku yang tidak memuat estimator dipindah ke ruas kanan sebagai berikut :

$$\hat{\boldsymbol{\beta}} = (\mathbf{I}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{WX})^{-1} \left(\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{Wy}-\left(\mathbf{X}^{T}\mathbf{WX}\right)^{-1}\mathbf{X}^{T}\mathbf{WT}\left(\mathbf{T}^{T}\mathbf{WT}\right)^{-1}\mathbf{T}^{T}\mathbf{Wy}\right)$$

Jika didefinisikan $\mathbf{M} = (\mathbf{I} - \left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX})^{-1}$,maka diperoleh estimator $\hat{\boldsymbol{\beta}}$ untuk komponen parametrik

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \mathbf{M}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy} - \mathbf{M}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{Wy}\\ &= \mathbf{M}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\left\{\mathbf{X}^T - \mathbf{X}^T\mathbf{WT}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\right\}\mathbf{Wy} \\ &= \mathbf{A(K)y}\end{aligned} \tag{4.9}$$

dimana $\mathbf{A(K)} = \mathbf{M}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\left\{\mathbf{X}^T - \mathbf{X}^T\mathbf{WT}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\right\}\mathbf{W}$.

Untuk mendapatkan $\hat{\boldsymbol{\eta}}$ yang bebas parameter, substitusi persamaan (4.7) ke persamaan (4.8)

$$\begin{aligned}\hat{\boldsymbol{\eta}} &= \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\left[\mathbf{T}^T\mathbf{Wy} - \mathbf{T}^T\mathbf{WX}\left\{\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\left\{\mathbf{X}^T\mathbf{Wy} - \mathbf{X}^T\mathbf{WT}\hat{\boldsymbol{\eta}}\right\}\right\}\right]\\ &= \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\left[\mathbf{T}^T\mathbf{Wy} - \mathbf{T}^T\mathbf{WX}\left\{\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy} - \left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT}\hat{\boldsymbol{\eta}}\right\}\right]\\ &= \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\left[\mathbf{T}^T\mathbf{Wy} - \mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy} + \mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT}\hat{\boldsymbol{\eta}}\right]\\ &= \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{Wy} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy} + \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\\ &\quad \left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT}\hat{\boldsymbol{\eta}}\end{aligned}$$

Selanjutnya suku yang memuat parameter dikelompokkan dalam satu ruas sebagai berikut :

$$\begin{aligned}\hat{\boldsymbol{\eta}} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT}\hat{\boldsymbol{\eta}} &= \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{Wy} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\\ &\quad \left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy}\end{aligned}$$

Persamaan di atas dapat diuraikan menjadi

$$\begin{aligned}\hat{\boldsymbol{\eta}}(\mathbf{J} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT} &= \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{Wy} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\\ &\quad \left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy}\end{aligned}$$

Suku yang tidak memuat estimator dipindah ke ruas kanan sebagai berikut :

$$\begin{aligned}\hat{\boldsymbol{\eta}} &= (\mathbf{I} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT})^{-1}\\ &\quad \left(\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{Wy} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy}\right)\end{aligned}$$

Jika didefinisikan $\mathbf{N} = (\mathbf{I} - \left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{WT})^{-1}$, maka diperoleh estimator $\hat{\boldsymbol{\eta}}$ untuk komponen nonparametrik

$$\begin{aligned}\hat{\boldsymbol{\eta}} &= \mathbf{N}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{Wy} - \mathbf{N}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\mathbf{Wy}\\ &= \mathbf{N}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\left\{\mathbf{T}^T - \mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\right\}\mathbf{Wy}\\ &= \mathbf{B(K)y}\end{aligned} \tag{4.10}$$

dimana $\mathbf{B(K)} = \mathbf{N}\left(\mathbf{T}^T\mathbf{WT}\right)^{-1}\left\{\mathbf{T}^T - \mathbf{T}^T\mathbf{WX}\left(\mathbf{X}^T\mathbf{WX}\right)^{-1}\mathbf{X}^T\right\}\mathbf{W}$.

Setelah mendapatkan estimator untuk komponen parametrik dan nonparametrik, berikutnya menentukan estimator model regresi semiparametrik birespon dengan pendekatan Deret Fourier.

$$\begin{aligned}\hat{\mathbf{y}} &= \hat{\mathbf{f}}(\mathbf{x},\mathbf{t}) \\ &= \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{T}\hat{\boldsymbol{\eta}} \\ &= \mathbf{XA(K)y} + \mathbf{TB(K)y} \\ &= (\mathbf{XA(K)} + \mathbf{TB(K)})\mathbf{y} \\ &= \mathbf{C(K)y}\end{aligned}$$

dengan $\mathbf{C(K)} = \mathbf{XA(K)} + \mathbf{TB(K)}$ adalah matriks berukuran $2n \times 2n$. Estimator model regresi semiparametrik birespon dengan pendekatan Deret Fourier terdiri atas estimator komponen parametrik persamaan (4.9) dan estimator komponen nonparametrik Deret Fourier pada persamaan (4.10). $\mathbf{A}(\mathbf{K})$ adalah matriks heat untuk komponen parametrik. $\mathbf{B}(\mathbf{K})$ matriks heat untuk komponen nonparametrik. Matriks heat untuk model regresi semiparametrik birespon dengan pendekatan Deret Fourier disimbolkan dengan $\mathbf{C}(\mathbf{K})$. $\mathbf{K}= (K_1, \dots, K_r)$ menunjukkan parameter osilasi yang termuat dalam matriks $\mathbf{C}(\mathbf{K})$. Nilai K optimal pada model diperoleh dengan menggunakan metode GCV (*Generalized Cross Validation*) yang formulanya diberikan sebagai berikut :

$$GCV\ (K_1, \dots, K_r) = \frac{MSE\ (K_1, \dots, K_r)}{((2n)^{-1} trace(\boldsymbol{I} - \boldsymbol{C}(K_1, \dots, K_r)))^2} \tag{4.11}$$

dengan

$$MSE\ (K_1, \dots, K_r) = (2n)^{-1} \boldsymbol{y}^{\boldsymbol{T}} \big(\boldsymbol{I} - \boldsymbol{C}(K_1, \dots, K_r)\big)^T \big(\boldsymbol{I} - \boldsymbol{C}(K_1, \dots, K_r)\big)\boldsymbol{y}\ .$$

r menunjukkan banyaknya variabel komponen nonparametrik.

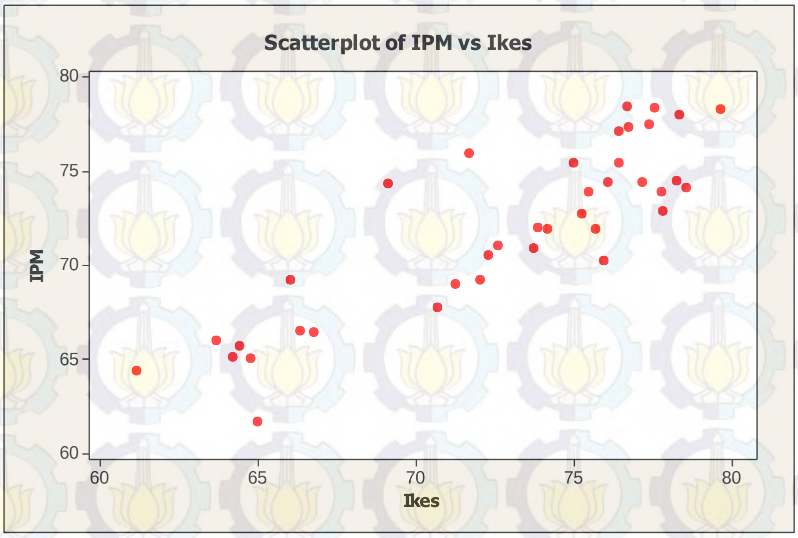
4.2 Model Regresi Semiparametrik Birespon dengan Pendekatan Deret Fourier untuk IPM dan IMH.

Pola hubungan antara IPM dan IMH dengan variabel Indeks Kesehatan Masyarakat, Indeks Pendidikan, PDRB (Prooduk Domestik Regional Bruto), Indeks PPP (*Purchasing Power Parity*), TPAK (Tingkat Partisipasi Angkatan Kerja), dan Persentase Kemiskinan di Jawa Timur dianalisis dengan menggunakan regresi. Sebelum dilakukan pemodelan regresi semiparametrik birespon dengan mengunakan pendekatan Deret Fourier, IPM dan IMH dimodelkan secara parametrik. Kelemahan pemodelan regresi parametrik birespon dalam hal pelanggaran asumsi dan signifikansi tiap prediktor dijadikan alasan untuk menggunaaan model regresi semiparametrik birespon dengan pendekatan Deret Fourier. Dalam pembahasan ini terlebih dahulu dilakukan pengujian dependensi antar variabel respon setelah deskripsi data digambarkan.

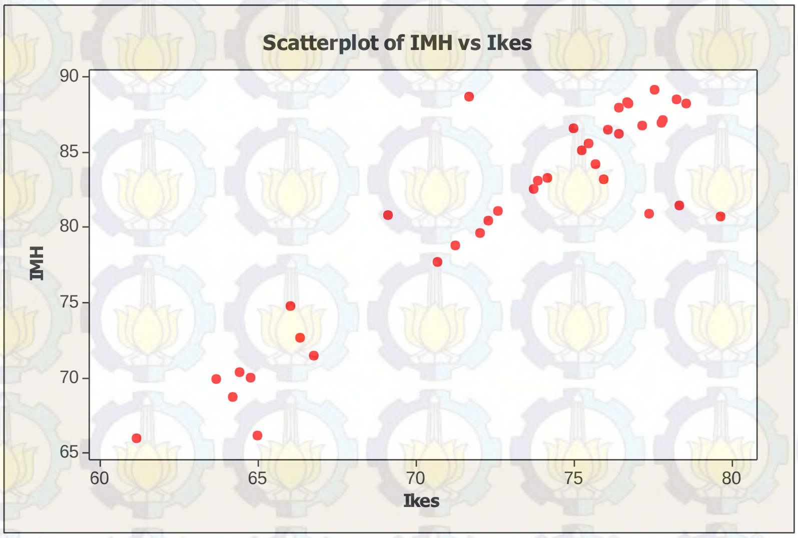
4.2.1 Deskripsi Data

Deskripsi data berguna untuk mengetahui gambaran umum data tentang variabel respon dan prediktor yang digunakan. Gambaran umum yang dibahas adalah daerah maksimum, daerah minimum dan hubungan antar respon dan prediktor dalam *scatter plot*.

Pada tahun 2012 nilai IPM untuk kabupaten dan kota di Jawa Timur yang tertinggi sebesar 78,43% dicapai Kota Malang. Kabupaten Sampang mempunyai nilai IPM terendah di Jawa Timur untuk tahun 2012 sebesar 61,67%. Nilai IMH untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kota Surabaya dengan nilai 89,11%. Nilai IMH yang terendah diperoleh Kabupaten Probolinggo sebesar 65,91%.



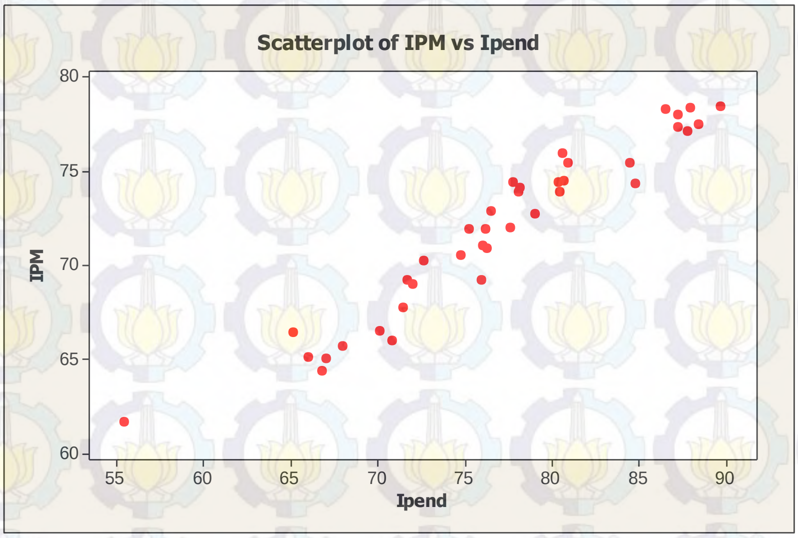
Gambar 4.1 *Scatter Plot* antara Indeks Kesehatan dengan IPM



Gambar 4.2 *Scatter Plot* antara Indeks Kesehatan dengan IMH

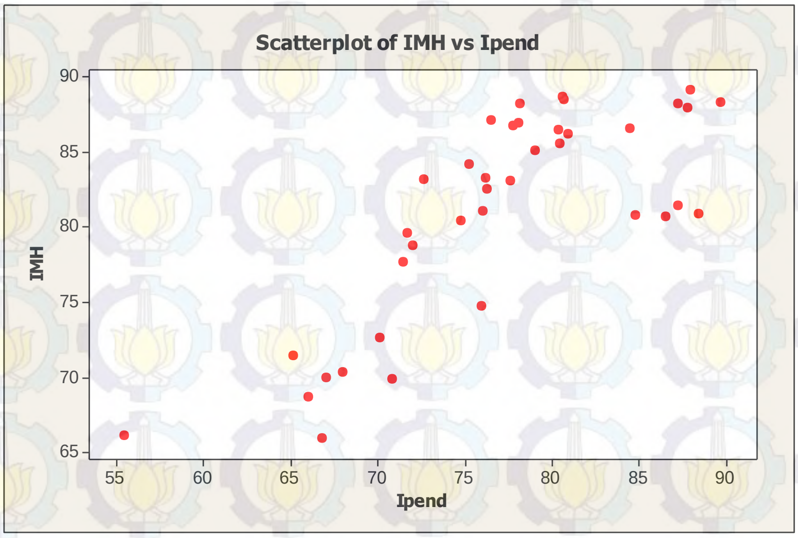
Prediktor indeks kesehatan untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kota Blitar dengan nilai 79,67%. Nilai indeks kesehatan yang terendah diperoleh Kabupaten Probolinggo sebesar 61,17%. Hubungan antara indeks kesehatan dan IPM dapat dilihat dari *scatter plot* Gambar 4.1. Berdasarkan Gambar 4.1 tampak bahwa jika indeks kesehatan meningkat maka IPM cenderung meningkat. Hubungan peningkatan ini terlihat linier. Hubungan antara indeks kesehatan dan IMH dapat dilihat dari *scatter plot* Gambar 4.2. Berdasarkan Gambar 4.2 tampak bahwa jika indeks kesehatan meningkat maka IMH cenderung meningkat. Hubungan peningkatan ini terlihat linier.

Prediktor indeks pendidikan untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kota Malang dengan nilai 89,72%. Nilai indeks pendidikan yang terendah diperoleh Kabupaten Sampang sebesar 55,46%. Hubungan antara indeks pendidikan dan IPM dapat dilihat dari *scatter plot* Gambar 4.3. Berdasarkan Gambar 4.3 tampak bahwa jika indeks pendidikan meningkat maka IPM cenderung meningkat. Hubungan peningkatan ini terlihat linier.



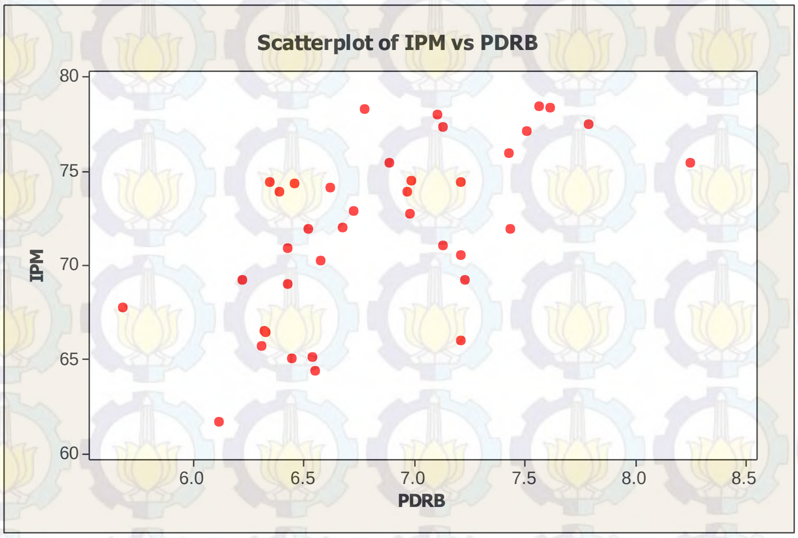
Gambar 4.3 *Scatter Plot* antara Indeks Pendidikan dengan IPM

Hubungan antara indeks pendidikan dan IMH dapat dilihat dari *scatter plot* Gambar 4.4. Berdasarkan Gambar 4.4 tampak bahwa jika indeks pendidikan meningkat maka IMH cenderung meningkat. Hubungan peningkatan ini terlihat linier.



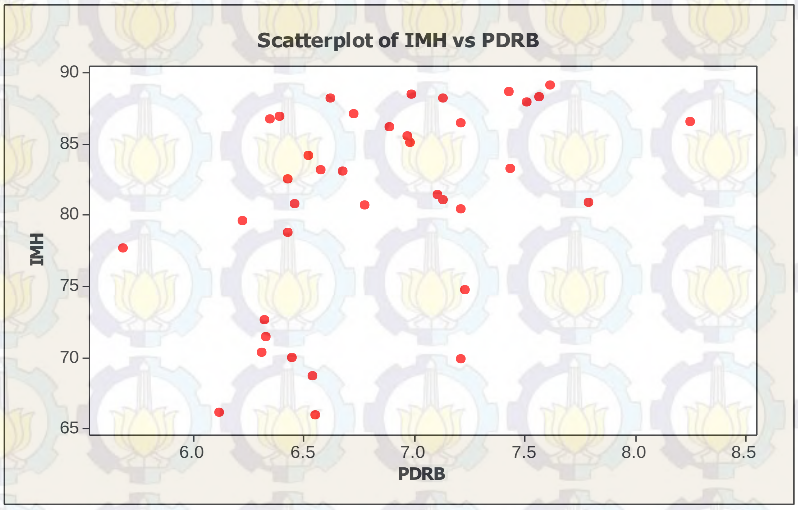
Gambar 4.4 *Scatter Plot* antara Indeks Pendidikan dengan IMH

Prediktor PDRB untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kota Batu dengan nilai 8,25%. Nilai PDRB yang terendah diperoleh Kabupaten Bojonegoro sebesar 5,68%.



Gambar 4.5 *Scatter Plot* antara PDRB dengan IPM

Hubungan antara PDRB dan IPM dapat dilihat dari *scatter plot* Gambar 4.5. Berdasarkan Gambar 4.5 tampak bahwa jika PDRB cenderung meningkat maka IPM juga terlihat cenderung meningkat. Hubungan peningkatan ini tidak mengikuti pola tertentu. Perlu dilakukan uji linieritas untuk mengetahui apakah pola data demikian masih dapat didekati dengan pola linier atau tidak.

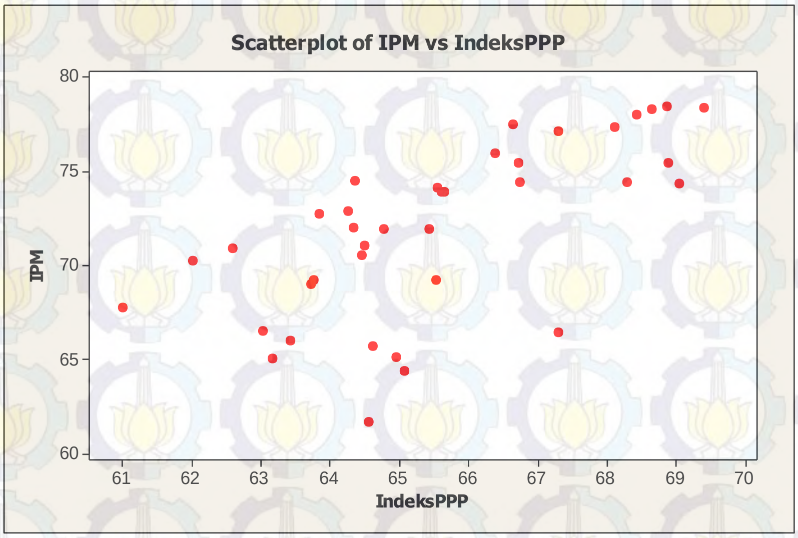


Gambar 4.6 *Scatter Plot* antara PDRB dengan IMH

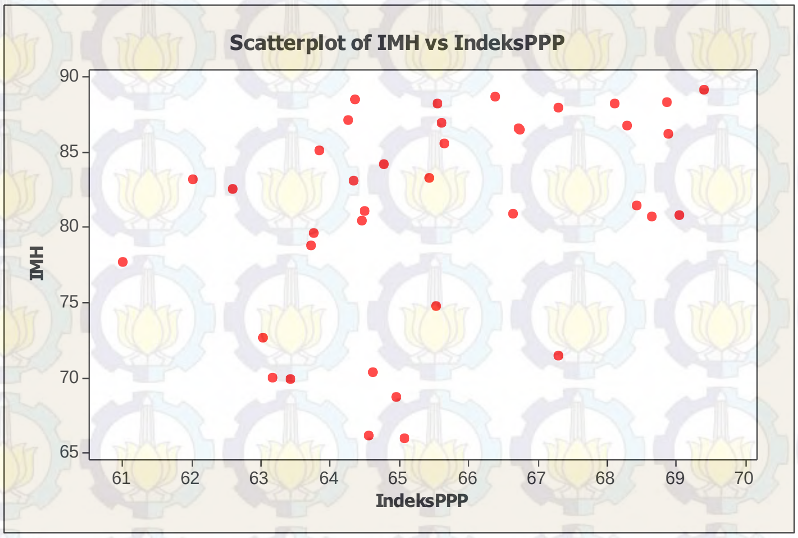
Untuk hubungan PDRB dan IMH dapat dilihat dari *scatter plot* Gambar 4.6. Berdasarkan Gambar 4.6 tampak bahwa jika PDRB cenderung meningkat maka

IMH juga terlihat cenderung meningkat. Hubungan peningkatan ini tidak mengikuti pola tertentu. Perlu dilakukan uji linieritas untuk mengetahui apakah pola data demikian masih dapat didekati dengan pola linier atau tidak.

Prediktor Indeks PPP untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kota Surabaya dengan nilai 69,42%. Nilai indeks PPP yang terendah diperoleh Kabupaten Bojonegoro sebesar 61,02%. Hubungan antara indeks PPP dan IPM dapat dilihat dari *scatter plot* Gambar 4.7. Berdasarkan Gambar 4.7 tampak bahwa jika indeks PPP cenderung meningkat maka IPM juga terlihat cenderung meningkat secara umum. Hubungan peningkatan ini tidak mengikuti pola tertentu. Perlu dilakukan uji linieritas untuk mengetahui apakah pola data demikian masih dapat didekati dengan pola linier atau tidak. Pola hubungan antara indeks PPP dan IMH dapat dilihat dari *scatter plot* Gambar 4.8. Berdasarkan Gambar 4.8 tampak bahwa jika indeks PPP cenderung meningkat maka IMH juga terlihat cenderung meningkat secara umum. Hubungan peningkatan ini tidak mengikuti pola tertentu. Perlu dilakukan uji linieritas untuk mengetahui apakah pola data demikian masih dapat didekati dengan pola linier atau tidak.

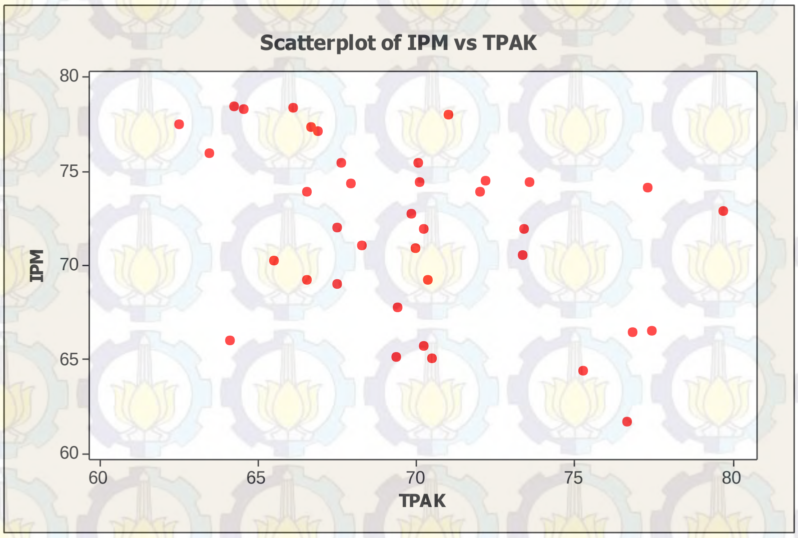


Gambar 4.7 *Scatter Plot* antara Indeks PPP dengan IPM



Gambar 4.8 *Scatter Plot* antara Indeks PPP dengan IMH

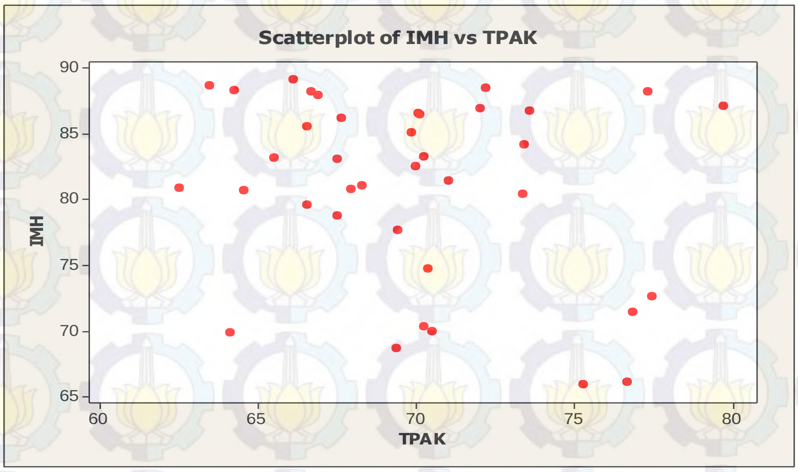
Prediktor TPAK untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kabupaten Pacitan dengan nilai 79,43%. Nilai TPAK yang terendah diperoleh Kota Madiun sebesar 62,53%.



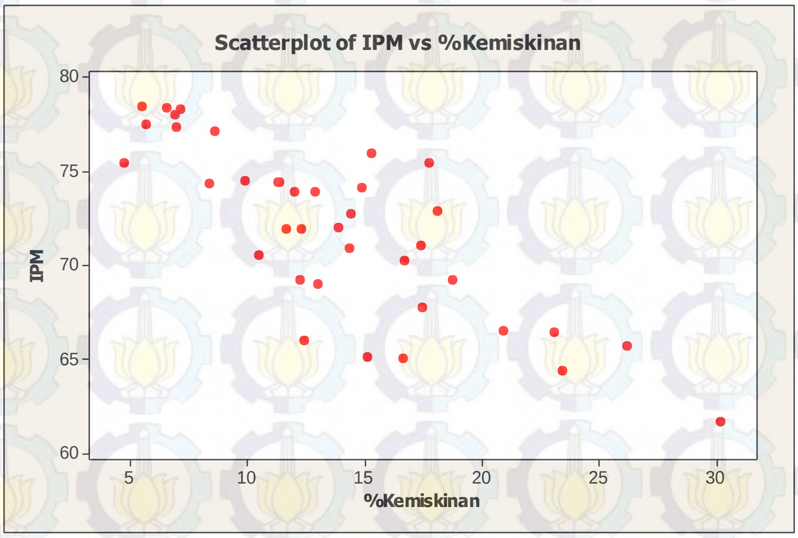
Gambar 4.9 *Scatter Plot* antara TPAK dengan IPM

Hubungan antara TPAK dan IPM dapat dilihat dari *scatter plot* Gambar 4.9. Berdasarkan Gambar 4.9 tampak bahwa tidak ada pola tertentu antara TPAK dengan IPM, namun jika TPAK cenderung meningkat maka IPM juga terlihat cenderung turun nilainya, TPAK yang cenderung lebih rendah nilainya, maka

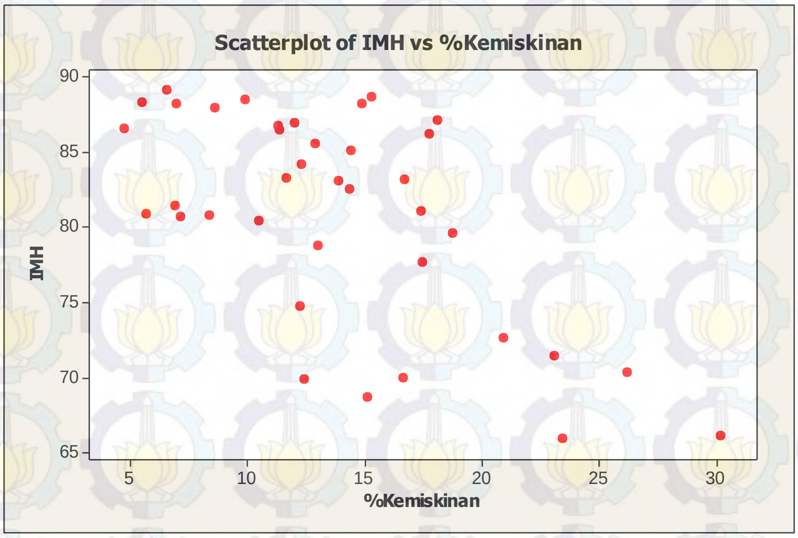
IPM cenderung lebih tinggi. Hubungan antara TPAK dan IMH dapat dilihat dari *scatter plot* Gambar 4.10. Berdasarkan Gambar 4.10 bahwa tidak ada pola tertentu antara TPAK dengan IMH, namun tampak bahwa jika TPAK cenderung meningkat maka IMH juga terlihat cenderung turun nilainya, TPAK yang cenderung lebih rendah nilainya, maka IMH cenderung lebih tinggi.



Gambar 4.10 *Scatter Plot* antara TPAK dengan IMH



Gambar 4.11 *Scatter Plot* antara Persentase Kemiskinan dengan IPM



Gambar 4.12 *Scatter Plot* antara Persentase Kemiskinan dengan IMH

Prediktor persentase kemiskinan untuk kabupaten dan kota di Jawa Timur yang tertinggi pada tahun 2012 dicapai oleh Kabupaten Sampang dengan nilai 27,87%. Nilai persentase kemiskinan terendah diperoleh Kota Batu sebesar 4.45%. Hubungan antara persentase kemiskinan dan IPM dapat dilihat dari *scatter plot* Gambar 4.11. Berdasarkan Gambar 4.11 tampak bahwa jika persentase kemiskinan cenderung meningkat maka IPM juga terlihat cenderung turun nilainya. Hubungan antara persentase kemiskinan dan IMH dapat dilihat dari *scatter plot* Gambar 4.12. Berdasarkan Gambar 4.12 bahwa jika persentase kemiskinan cenderung meningkat maka IMH juga terlihat cenderung turun nilainya.

4.2.2 Pengujian Dependensi Variabel Respon

Pengujian dependensi antar variabel respon digunakan uji *Bartlett Sphericity* pada data IPM dan IMH. Dalam pengujian ini diharapkan antar variabel berkorelasi.

Hipotesis :

H_0 : Antar variabel respon independen

H_1 : Antar variabel respon dependen

Diperoleh $\chi^2_{hitung} = 50{,}4233 > \chi^2_{tabel} = 3{,}842$ maka tolak H_0 sehingga antar variabel respon dependen. Antar variabel respon memiliki korelasi 87,08%.

4.2.3 Kelemahan Model Regresi Parametrik Birespon untuk IPM dan IMH

Dalam model regresi parametrik terdapat asumsi – asumsi residual yang harus dipenuhi yaitu asumsi multikolinieritas, identik, independen, dan berdistribusi normal multivariat. Jika tidak memenuhi salah satu dari asumsi tersebut maka harus diatasi. Dalam mengatasi pelanggaran asumsi terdapat persoalan, misalnya residual dari model regresi yang baru dilakukan penanganan terdapat kasus pelanggaran asumsi lagi. Selain itu kelemahan regresi parametrik yang dilihat dalam pemodelan IPM dan IMH ini adalah signifikansi parameter. Regresi parametrik birespon untuk IPM dan IMH adalah sebagai berikut :

a. Pengujian Normal Multivariat Variabel Respon

Pengujian normal multivariat dilakukan untuk mengetahui bahwa variabel IPM dan IMH berdistribusi normal multivariat.

Hipotesis :

H_0 : Variabel respon berdistribusi normal multivariat

H_1 : Variabel respon tidak berdistribusi normal multivariat

Diperoleh hasil bahwa t = 52,6316 % yang berada di sekitar 50 %, sehingga diputuskan untuk gagal tolak H_0. Artinya variabel respon berdistribusi normal multivariat.

b. Model Regresi Parametrik Birespon

Diberikan model linier regresi parametrik birespon sebagai berikut

$$\hat{y}_{1i} = \hat{\beta}_{01} + \hat{\beta}_{11}X_{1i} + \hat{\beta}_{21}X_{2i} + \cdots + \hat{\beta}_{61}X_{6i}$$

$$\hat{y}_{2i} = \hat{\beta}_{02} + \hat{\beta}_{12}X_{1i} + \hat{\beta}_{22}X_{2i} + \cdots + \hat{\beta}_{62}X_{6i}$$

Estimasi kurva regresi parametrik birespon untuk model IPM dan IMH di Jawa Timur diberikan oleh

$$\hat{y}_{1i} = -1{,}1476 + 0{,}3287X_{1i} + 0{,}354X_{2i} + 0{,}25X_{3i} + 0{,}332X_{4i} - 0{,}0341X_{5i} + 0{,}052X_{6i}$$

$$\hat{y}_{2i} = -8{,}8827 + 0{,}946X_{i1} + 0{,}376X_{2i} + 1{,}232X_{3i} - 0{,}357X_{4i} + 0{,}063X_{5i} + 0{,}1835X_{6i}$$

dimana y_1: IPM, y_2: IMH, X_1 : Indeks Kesehatan, X_2: Indeks Pendidikan, X_3: PDRB, X_4: Indeks PPP, X_5: TPAK, X_6 : Persentase Kemiskinan.

Model tersebut diselidiki apakah telah signifikan secara serentak dan individu, serta memenuhi asumsi residual identik, independen dan berdistribusi normal multivariat.

c. Pengujian Signifikansi Serentak

Uji serentak adalah pengujian secara bersama-sama terhadap parameter yang terdapat di dalam persamaan regresi. Uji *Wilks Lambda* digunakan untuk pengujian serentak.

Hipotesis :

$H_0 : \beta_{11} = \beta_{12} \dots = \beta_{62} = 0$ (model tidak signifikan)

H_1: paling sedikit ada satu $\beta_{pq} \neq 0$

dimana $p = 1,2,3,4,5,6$ dan $q = 1,2$ (model signifikan)

Dengan menggunakan persamaan (2.4) diperoleh:

$\Lambda_{hitung} = 0{,}0000010151 < \Lambda_{\text{tabel}} = \Lambda_{0.05,2,6,31} = 0{,}4746.$

Oleh sebab itu H_0 ditolak dimana secara keseluruhan parameter tidak sama dengan nol sehingga model signifikan.

d. Pengujian Signifikansi Individu

Tabel 4.1 Tabel Uji Individu

Variabel Respon	Parameter	$\hat{\beta}$	$\widehat{SE(\beta)}$	t_{hitung}	Kesimpulan
Y_1	Intercept	-0,8639	3,7286	-0,3078	Tidak Signifikan
	X_1	0,3287	0,0278	11,8338	Signifikan
	X_2	0,3541	0,0328	10,8057	Signifikan
	X_3	0,25	0,2209	1,1320	Tidak Signifikan
	X_4	0,3321	0,0532	6,2395	Signifikan
	X_5	-0,0341	0,0245	-1,3903	Tidak Signifikan
	X_6	0,052	0,0293	1,7766	Tidak Signifikan
Y_2	Intercept	-8,8827	22,1937	-0,4002	Tidak Signifikan
	X_1	0,9463	0,1653	5,7241	Signifikan
	X_2	0,3763	0,1950	1,9293	Tidak Signifikan
	X_3	1,2315	1,3147	0,9367	Tidak Signifikan
	X_4	-0,3570	0,3168	-1,1268	Tidak Signifikan
	X_5	0,0631	0,1458	0,4329	Tidak Signifikan
	X_6	0,1835	0,1742	1,0532	Tidak Signifikan

Pengujian ini bertujuan untuk melihat pengaruh setiap variabel prediktor terhadap variabel-variabel respon

Hipotesis :

$H_0 : \beta_{lj} = 0$ (parameter model prediktor ke– l ; $l = 1,2,3,4,5,6$ terhadap respon ke – j ; $j = 1,2$ tidak berpengaruh signifikan)

$H_1 : \beta_{lj} \neq 0$ (parameter model prediktor ke– ; $l = 1,2,3,4,5,6$, terhadap respon ke – j ; $j = 1,2$ berpengaruh signifikan)

Tabel 4.1 berikut adalah tabel uji parsial dengan t tabel = $t_{0.025,35} = 2.342$. Berdasarkan Tabel 4.1, dapat dilihat bahwa banyak parameter yang tidak signifikan, sehingga H_0 gagal ditolak. Untuk respon IPM, hanya variabel indeks kesehatan dan indeks pendidikan yang signifikan. Untuk respon IMH, hanya variabel indeks kesehatan yang signifikan. Model regresi parametrik birespon ini hanya mengakomodasi sedikit prediktor saja.

e. Pengujian Asumsi Residual Identik

Untuk menguji apakah residual data identik, maka asumsi yang harus dipenuhi dalam melakukan pemodelan regresi multivariat adalah residual memiliki matriks varian-kovarian ($\mathbf{\Sigma}$) yang homogen. Untuk menguji syarat ini dapat dipergunakan statistik uji *Box's M*.

Hipotesis :

H_0 : $\mathbf{\Sigma_1} = \mathbf{\Sigma_2} = \mathbf{\Sigma}$

H_1 : minimal ada satu $\mathbf{\Sigma_i} \neq \mathbf{\Sigma_j}$ untuk $i \neq j, i = 1,2, j = 1,2.$

Hasil dari statistik uji ini adalah sebagai berikut :

Tabel 4.2 Uji Box's M

Box's M	46,006
F	14,415
df1	3
df2	23.330
Sig.	0,000

Berdasarkan Tabel 4.2 diperoleh p-value < α=0,05 sehingga H_0 ditolak yang berarti matriks varian-kovarian residual adalah heterogen dan dapat disimpulkan residual tidak identik atau terjadi heteroskedastisitas.

f. Pengujian Asumsi Residual Independen

Untuk menguji apakah residual data independen digunakan uji *Bartlett Sphericity*.

Hipotesis :

H_0 : Residual independen

H_1 : Residual tidak independen

Dengan menggunakan persaamaan (2.7), diperoleh $\chi^2_{hitung} = 9{,}3663 > \chi^2_{tabel} = 3{,}842$ maka tolak H_0 sehingga residual tidak independen.

g. Pengujian Asumsi Residual Berdistribusi Normal Multivariat

Pengujian normal multivariat dilakukan untuk mengetahui bahwa residual berdistribusi normal multivariat.

Hipotesis :

H_0 : Residual berdistribusi normal multivariat

H_1 : Residual tidak berdistribusi normal multivariat

Berdasarkan persamaan (2.8), diperoleh hasil bahwa t = 39,4737 % yang tidak berada di sekitar 50 %, sehingga diputuskan untuk tolak H_0. Artinya residual tidak berdistribusi normal multivariat.

Dengan menggunakan regresi parametrik birespon, tidak satupun asumsi yang dipenuhi dan banyak variabel yang secara parsial tidak signifikan sehingga perlu dicari model regresi yang lain untuk memodelkan variabel Indeks Kesehatan, Indeks Pendidikan, PDRB, Indeks PPP, TPAK, dan Persentase Kemiskinan terhadap IPM dan IMH. Metode lain yang disarankan adalah regresi semiparametrik birespon dengan pendekatan Deret Fourier yang bebas dari asumsi.

4.2.4 Identifikasi Variabel Komponen Parametrik dan Nonparametrik

Dalam melakukan identifikasi variabel komponen parametrik dan nonparametrik dilakukan dengan menggunakan interpretasi *scatter plot* dan uji linieritas. Berdasarkan *scatter plot* pada pembahasan sebelumnya data berpasangan yang memiliki hubungan peningkatan linier, prediktor penyusunnya digolongkan dalam variabel komponen parametrik, sedangkan data berpasangan yang tidak berpola, prediktor penyusunnya termasuk dalam variabel komponen nonparametrik. Hasil identifikasi *scatter plot* disajikan dalam Tabel 4.3.

Tabel 4.3 Tabel Hasil Identifikasi *Scatter Plot*

Respon	Prediktor	Hasil Identifikasi
IPM	Indeks Kesehatan	Parametrik Linier
IPM	Indeks Pendidikan	Parametrik Linier
IPM	PDRB	Nonparametrik
IPM	IndeksPPP	Nonparametrik
IPM	TPAK	Nonparametrik
IPM	Persentase Kemiskinan	Nonparametrik
IMH	Indeks Kesehatan	Parametrik Linier
IMH	Indeks Pendidikan	Parametrik Linier
IMH	PDRB	Nonparametrik
IMH	IndeksPPP	Nonparametrik
IMH	TPAK	Nonparametrik
IMH	Persentase Kemiskinan	Nonparametrik

Berdasarkan Tabel 4.3, prediktor Indeks Kesehatan, dan Indeks Pendidikan, berpola linier terhadap IPM dan IMH, sehingga kedua variabel tersebut digunakan sebagai komponen parametrik x_1, dan x_2. Variabel PDRB, Indeks PPP, TPAK dan Persentase Kemiskinan tidak mengikuti pola tertentu terhadap IPM dan IMH, sehingga menjadi komponen nonparametrik masing–masing t_1 , t_2 , t_3 dan t_4.

Untuk pengujian linieritas, hipotesis yang diajukan adalah :

H_0 : terdapat hubungan linier antara prediktor terhadap respon.

H_1 : tidak terdapat hubungan linier antara prediktor terhadap respon.

Kriteria pengambilan keputusan, jika *p-value* < 0,05 maka tolak H_0 dan sebaliknya jika *p-value* > 0,05 berarti gagal tolak H_0. Hasil uji linieritas disajikan dalam Tabel 4.4 sebagai berikut :

Tabel 4.4 Tabel Hasil Uji Linieritas

Respon	Prediktor	Perbandingan p-value dan α	Kesimpulan
IPM	Indeks Kesehatan	0,393>0,05	Linier
IPM	Indeks Pendidikan	0,292>0,05	Linier
IPM	PDRB	0,613>0,05	Linier
IPM	Indeks PPP	0,976>0,05	Linier
IPM	TPAK	0,011<0,05	Non Linier
IPM	Persentase Kemiskinan	0,015<0,05	Non Linier
IMH	Indeks Kesehatan	0,307>0,05	Linier
IMH	Indeks Pendidikan	0,723>0,05	Linier
IMH	PDRB	0,624>0,05	Linier
IMH	Indeks PPP	0,924>0,05	Linier
IMH	TPAK	0,005<0,05	Non Linier
IMH	Persentase Kemiskinan	0,013<0,05	Non Linier

Berdasarkan Tabel 4.4, prediktor TPAK dan Tingkat Kemiskinan tidak linier terhadap IPM dan IMH, sehingga kedua variabel tersebut digunakan sebagai komponen nonparametrik masing – masing t_1 dan t_2, sedangkan variabel Indeks Kesehatan, Indeks Pendidikan, PDRB, Indeks PPP, linier terhadap IPM dan IMH, sehingga menjadi komponen parametrik masing–masing x_1 , x_2, x_3 dan x_4. Model regresi semiparametrik untuk IPM dan IMH yang terbaik diperoleh dengan membandingkan model regresi semiparametrik yang variabel prediktor parametrik dan nonparametriknya dipilih berdasarkan *scatter plot* (p =2, r = 4) dan pengujian linieritas (p =4, r =2).

4.2.5. Model Regresi Semiparametrik IPM dan IMH untuk p =2, r = 4

Pada *scatter plot* dapat dilihat bahwa terdapat kecenderungan pengulangan nilai respon pada nilai prediktor yang berbeda. Untuk variabel PDRB dan Indeks PPP, terdapat kecenderungan bahwa terjadi peningkatan jika nilai respon IPM dan IMH cenderung lebih besar. Sebaliknya pada variabel TPAK, dan persentase kemiskinan, terdapat kecenderungan bahwa terjadi penurunan jika nilai respon

IPM dan IMH lebih besar. Hal tersebut mendasari digunakan pendekatan Deret Fourier untuk variabel-variabel komponen nonparametrik tersebut. Oleh sebab itu indeks kesehatan (x_1) dan indeks pendidikan (x_2) ditetapkan sebagai variabel komponen parametrik linier. Variabel PDRB (t_1), Indeks PPP (t_2), TPAK (t_3), dan persentase kemiskinan (t_4) sebagai variabel komponen nonparametrik.

Diberikan data berpasangan $(x_1, x_2, t_1, t_2, t_3, t_4,\ y_1, y_2)$. Variabel prediktor x_1, dan x_2 berpola linier, variabel prediktor $t_1, t_2, t_3,$ dan t_4 tidak mengikuti pola tertentu. Model regresi semiparametrik birepon yang memuat variabel tersebut dinyatakan sebagai berikut.

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \beta_{21}x_{2i} + f_1(t_{1i}) + f_1(t_{2i}) + f_1(t_{3i}) + f_1(t_{4i}) + \varepsilon_{1i}$$

$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \beta_{22}x_{2i} + f_2(t_{1i}) + f_2(t_{2i}) + f_2(t_{3i}) + f_2(t_{4i}) + \varepsilon_{2i}$$

Jika fungsi $f_j(t_{ri})$, $j = 1,2$ dan $r = 1,2,3,4$ didekati dengan Deret Fourier maka bentuk model menjadi.

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \beta_{21}x_{2i} + \alpha_{01}^{*} + \gamma_{11}t_{1i} + \sum_{k=1}^{K}(\alpha_{k11}\cos kt_{1i}) + \gamma_{12}t_{2i} +$$
$$\sum_{k=1}^{K}(\alpha_{k12}\cos kt_{2i}) + \gamma_{13}t_{3i} + \sum_{k=1}^{K}(\alpha_{k13}\cos kt_{3i}) + \gamma_{14}t_{4i} + \sum_{k=1}^{K}(\alpha_{k14}\cos kt_{4i}) + \varepsilon_{1i}$$

$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \beta_{22}x_{2i} + \alpha_{02}^{*} + \gamma_{21}t_{1i} + \sum_{k=1}^{K}(\alpha_{k21}\cos kt_{1i}) + \gamma_{22}t_{2i} +$$
$$\sum_{k=1}^{K}(\alpha_{k22}\cos kt_{2i}) + \gamma_{23}t_{3i} + \sum_{k=1}^{K}(\alpha_{k23}\cos kt_{3i}) + \gamma_{24}t_{4i} + \sum_{k=1}^{K}(\alpha_{k24}\cos kt_{4i}) + \varepsilon_{2i}$$

Model regresi semiparametrik birespon dengan pendekatan Deret Fourier di atas bergantung pada parameter osilasi K.

Untuk mendapatkan estimator yang terbaik dalam model regresi semiparametrik birespon dengan pendekatan Deret Fourier dilakukan pemilihan K optimal dengan menggunakan metode GCV. Nilai K dipilih berdasarkan nilai GCV minimum. Nilai K merupakan jumlah osilasi gelombang cosinus pada model regresi dengan pendekatan Deret Fourier. Nilai K yang semakin besar mengakibatkan model semakin kompleks dan osilasi dari kurva regresi semakin rapat. Untuk itu nilai K dibatasi sampai $K = 3$. Rumus untuk menghitung nilai GCV telah diberikan dalam persamaan (4.11). Tabel 4.5 merupakan Tabel perolehan GCV untuk nilai $K = 1, 2, 3$ dengan $p = 2, r = 4$.

Tabel 4.5 Tabel Nilai GCV untuk p =2, r = 4

Nilai *K*	Nilai GCV
1	15.444.466
2	9.031.718
3	2.927.628

Dengan memperhatikan *K*=3, diperoleh GCV minimum jika dibandingkan dengan nilai GCV untuk *K*=1, dan *K*=2. Model regresi semiparametrik birespon dengan pendekatan Deret Fourier dengan *K*=3 dipilih untuk memodelkan IPM dan IMH di Jawa Timur. Estimasi model regresi semiparametrik birespon dengan pendekatan Deret Fourier untuk *K*=3 adalah sebagai berikut :

$$\hat{y}_{1i} = \hat{\beta}_{01} + \hat{\beta}_{11}x_{1i} + \hat{\beta}_{21}x_{2i} + \hat{\alpha}^*_{01} + \hat{\gamma}_{11}t_{1i} + \hat{\alpha}_{111}\cos t_{1i} + \hat{\alpha}_{211}\cos 2t_{1i} + \hat{\alpha}_{311}\cos 3t_{1i} +$$
$$\hat{\gamma}_{12}t_{2i} + \hat{\alpha}_{112}\cos t_{2i} + \hat{\alpha}_{212}\cos 2t_{2i} + \hat{\alpha}_{312}\cos 3t_{2i} + \hat{\gamma}_{13}t_{3i} + \hat{\alpha}_{113}\cos t_{3i} +$$
$$\hat{\alpha}_{213}\cos 2t_{3i} + \hat{\alpha}_{313}\cos 3t_{3i} + \hat{\gamma}_{14}t_{4i} + \hat{\alpha}_{114}\cos t_{4i} + \hat{\alpha}_{214}\cos 2t_{4i} + \hat{\alpha}_{314}\cos 3t_{4i}$$
$$\hat{y}_{2i} = \hat{\beta}_{02} + \hat{\beta}_{12}x_{1i} + \hat{\beta}_{22}x_{2i} + \hat{\alpha}^*_{02} + \hat{\gamma}_{21}t_{1i} + \hat{\alpha}_{121}\cos t_{1i} + \hat{\alpha}_{221}\cos 2t_{1i} + \hat{\alpha}_{321}\cos 3t_{1i} +$$
$$\hat{\gamma}_{22}t_{2i} + \hat{\alpha}_{122}\cos t_{2i} + \hat{\alpha}_{222}\cos 2t_{2i} + \hat{\alpha}_{322}\cos 3t_{2i} + \hat{\gamma}_{23}t_{3i} + \hat{\alpha}_{123}\cos t_{3i} +$$
$$\hat{\alpha}_{223}\cos 2t_{3i} + \hat{\alpha}_{323}\cos 3t_{3i} + \hat{\gamma}_{24}t_{4i} + \hat{\alpha}_{124}\cos t_{4i} + \hat{\alpha}_{224}\cos 2t_{4i} + \hat{\alpha}_{324}\cos 3t_{4i}$$

Nilai estimasi untuk masing –masing parameter regresi semiparametrik birespon dengan pendekatan Deret Fourier diperoleh dari pemrograman. Model regresi semiparametrik birespon dengan pendekatan Deret Fourier untuk IPM dan IMH di Jawa Timur dengan p =2, r = 4 adalah sebagai berikut :

$$\hat{y}_{1i} = 0{,}01 + 0{,}326x_{1i} + 0{,}354x_{2i} + 0{,}178t_{1i} + 0{,}335\cos t_{1i} - 0{,}043\cos 2t_{1i} +$$
$$0{,}053\cos 3t_{1i} - 0{,}12t_{2i} - 0{,}147\cos t_{2i} + 0{,}079\cos 2t_{2i} + 0{,}006\cos 3t_{2i} +$$
$$0{,}006t_{3i} + 0{,}006\cos t_{3i} + 0{,}006\cos 2t_{3i} + 0{,}006\cos 3t_{3i} + 0{,}006t_{4i} +$$
$$0{,}006\cos t_{4i} + 0{,}006\cos 2t_{4i} + 0{,}006\cos 3t_{4i}$$
$$\hat{y}_{2i} = 0{,}939x_{1i} + 0{,}373x_{2i} + 0{,}842t_{1i} - 0{,}31\cos t_{1i} - 0{,}005\cos 2t_{1i} +$$
$$0{,}226\cos 3t_{1i} - 0{,}636t_{2i} - 1{,}314\cos t_{2i} + 0{,}524\cos 2t_{2i}$$

Model tersebut memiliki nilai MSE = 11,451. Model regresi semiparametrik birespon dengan pendekatan Deret Fourier tersebut tidak memuat variabel prediktor TPAK (t_3) dan persentase kemisikinan (t_4) dalam IMH (y_2) karena nilai estimator dari parameter untuk variabel tersebut sangat kecil. Dengan demikian perlu dimodelkan lagi IPM dan IMH di Jawa Timur dengan identifikasi variabel prediktor parametrik dan nonparametrik untuk p =4, r = 2 berdasarkan uji linieritas.

4.2.6. Model Regresi Semiparametrik IPM dan IMH untuk p =4, r = 2

Berdasarkan uji linieritas, indeks kesehatan (x_1) indeks pendidikan (x_2), PDRB (x_3), dan indeks PPP (x_4), telah ditetapkan sebagai variabel komponen parametrik linier. Variabel TPAK (t_1), dan persentase kemiskinan (t_2) sebagai variabel komponen nonparametrik. Diberikan data berpasangan ($x_1, x_2, x_3, x_4, t_1, t_2, y_1, y_2$). Variabel prediktor $x_1, x_2, x_3,$ dan x_4 berpola linier, variabel prediktor t_1, dan t_2 tidak mengikuti pola tertentu. Model regresi semiparametrik birepon yang memuat variabel tersebut dinyatakan sebagai berikut.

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \beta_{21}x_{2i} + \beta_{31}x_{3i} + \beta_{41}x_{4i} + f_1(t_{1i}) + f_1(t_{2i}) + \varepsilon_{1i}$$

$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \beta_{22}x_{2i} + \beta_{32}x_{1i} + \beta_{42}x_{2i} + f_2(t_{1i}) + f_2(t_{2i}) + \varepsilon_{2i}$$

Jika fungsi $f_j(t_{ri})$, $j = 1,2$ dan $r = 1,2$ didekati dengan Deret Fourier maka bentuk model menjadi.

$$y_{1i} = \beta_{01} + \beta_{11}x_{1i} + \beta_{21}x_{2i} + \beta_{31}x_{3i} + \beta_{41}x_{4i} + \alpha_{01}^{*} + \gamma_{11}t_{1i} + \sum_{k=1}^{K}(\alpha_{k11}\cos kt_{1i}) + \gamma_{12}t_{2i} + \sum_{k=1}^{K}(\alpha_{k12}\cos kt_{2i}) + \varepsilon_{1i}$$

$$y_{2i} = \beta_{02} + \beta_{12}x_{1i} + \beta_{22}x_{2i} + \beta_{32}x_{3i} + \beta_{42}x_{4i} + \alpha_{02}^{*} + \gamma_{21}t_{1i} + \sum_{k=1}^{K}(\alpha_{k21}\cos kt_{1i}) + \gamma_{22}t_{2i} + \sum_{k=1}^{K}(\alpha_{k22}\cos kt_{2i}) + \varepsilon_{2i}$$

Model regresi semiparametrik birespon dengan pendekatan Deret Fourier di atas bergantung pada parameter osilasi K.

Untuk mendapatkan estimator yang terbaik dalam model regresi semiparametrik birespon dengan pendekatan Deret Fourier dilakukan pemilihan K optimal dengan menggunakan metode GCV. Nilai K dipilih berdasarkan nilai GCV minimum. Nilai K merupakan jumlah osilasi gelombang cosinus pada model regresi dengan pendekatan Deret Fourier. Nilai K yang semakin besar mengakibatkan model semakin kompleks dan osilasi dari kurva regresi semakin rapat. Untuk itu nilai K dibatasi sampai K = 3. Rumus untuk menghitung nilai

GCV telah diberikan dalam persamaan (4.11). Tabel 4.6 merupakan Tabel perolehan GCV untuk nilai K = 1, 2, 3.

Tabel 4.6. Tabel Nilai GCV untuk $p=4, r=2$

Nilai K	Nilai GCV
1	1.166.465
2	948.207
3	753.586

Dengan memperhatikan K=3, diperoleh GCV minimum jika dibandingkan dengan nilai GCV untuk K=1, dan K=2. Model regresi semiparametrik birespon dengan pendekatan Deret Fourier dengan K=3 dipilih untuk memodelkan IPM dan IMH di Jawa Timur. Estimasi model regresi semiparametrik birespon dengan pendekatan Deret Fourier untuk K=3 adalah sebagai berikut :

$$\hat{y}_{1i} = \hat{\beta}_{01} + \hat{\beta}_{11}x_{1i} + \hat{\beta}_{21}x_{2i} + \hat{\beta}_{31}x_{3i} + \hat{\beta}_{41}x_{4i} + \hat{\alpha}_{01}^{*} + \hat{\gamma}_{11}t_{1i} + \hat{\alpha}_{111}\cos t_{1i} + \hat{\alpha}_{211}\cos 2t_{1i} + \hat{\alpha}_{311}\cos 3t_{1i} + \hat{\gamma}_{12}t_{2i} + \hat{\alpha}_{112}\cos t_{2i} + \hat{\alpha}_{212}\cos 2t_{2i} + \hat{\alpha}_{312}\cos 3t_{2i}$$

$$\hat{y}_{2i} = \hat{\beta}_{02} + \hat{\beta}_{12}x_{1i} + \hat{\beta}_{22}x_{2i} + \hat{\beta}_{32}x_{3i} + \hat{\beta}_{42}x_{4i} + \hat{\alpha}_{02}^{*} + \hat{\gamma}_{21}t_{1i} + \hat{\alpha}_{121}\cos t_{1i} + \hat{\alpha}_{221}\cos 2t_{1i} + \hat{\alpha}_{321}\cos 3t_{1i} + \hat{\gamma}_{22}3t_{2i} + \hat{\alpha}_{122}\cos t_{2i} + \hat{\alpha}_{222}\cos 2t_{2i} + \hat{\alpha}_{322}\cos 3t_{2i}$$

Nilai estimasi untuk masing – masing parameter koefisien regresi semiparametrik birespon dengan pendekatan Deret Fourier diperoleh dari pemrograman. Nilai tersebut disajikan dalam Tabel 4.7. Berdasarkan estimator Tabel 4.7 diperoleh model regresi semiparametrik birespon dengan pendekatan Deret Fourier untuk IPM dan IMH di Jawa Timur dengan $p=4, r=2$ sebagai berikut :

$$\hat{y}_{1i} = 0{,}043 + 0{,}326x_{1i} + 0{,}354x_{2i} + 0{,}182x_{3i} + 0{,}337x_{4i} - 0{,}042t_{1i} + 0{,}053\cos t_{1i} - 0{,}119\cos 2t_{1i} - 0{,}146\cos 3t_{1i} + 0{,}08t_{2i} + 0{,}035\cos t_{2i} + 0{,}035\cos 2t_{2i} + 0{,}035\cos 3t_{2i}$$

$$\hat{y}_{2i} = -1{,}015 + 0{,}939x_{1i} + 0{,}373x_{2i} + 0{,}842x_{3i} - 0{,}31x_{4i} - 0{,}005t_{1i} + 0{,}226\cos t_{1i} - 0{,}636\cos 2t_{1i} - 1{,}314\cos 3t_{1i} + 0{,}525t_{2i} - 1{,}029\cos t_{2i} + -1{,}029\cos 2t_{2i} - 1{,}029\cos 3t_{2i}$$

Tabel 4.7. Tabel Nilai Parameter Koefisien Regresi $p = 4, r = 2$

Parameter Respon 1	Nilai	Parameter Respon 2	Nilai
β_{01}	0,008	β_{02}	0,014
β_{11}	0,326	β_{12}	0,939
β_{21}	0,354	β_{22}	0,373
β_{31}	0,182	β_{32}	0,842
β_{41}	0,337	β_{42}	-0,31
α_{01}^*	0,035	α_{02}^*	-1,029
γ_{11}	-0,042	γ_{21}	-0,005
α_{111}	0,053	α_{121}	0,226
α_{211}	-0,119	α_{221}	-0,636
α_{311}	-0,146	α_{321}	-1,314
γ_{12}	0,08	γ_{22}	0,525
α_{112}	0,035	α_{122}	-1,029
α_{212}	0,035	α_{222}	-1,029
α_{312}	0,035	α_{322}	-1,029

Model regresi semiparametrik birespon dengan pendekatan Deret Fourier tersebut telah mengakomodasi semua variabel. Model ini memiliki nilai MSE = 3,052 dan GCV minimum = 753.586. Nilai – nilai tersebut lebih kecil dibandingkan model pertama dengan identifikasi prediktor berdasarkan *scatter plot*. Koefisien determinasi atau R^2 yang dimiliki model ini tinggi yaitu sebesar 0,9198. Artinya keragaman nilai respon telah mampu dijelaskan oleh variabel prediktor sebesar 91,98%.

Berdasarkan analisis tersebut diperoleh model regresi semiparametrik untuk IPM dan IMH yang terbaik yaitu model dengan $p = 4, r = 2$. Model tersebut mengakomodasi semua prediktor, memiliki nilai GCV dan MSE yang lebih kecil dan R^2 = 91,98%. Dalam kasus ini identifikasi variabel prediktor komponen parametrik dan nonparametrik dengan *scatter plot* memberikan model yang kurang baik dibandingkan dengan identifikasi dengan uji linieritas. Namun perlu penelitian lebih lanjut untuk dapat menyimpulkan bahwa dalam identifikasi variabel prediktor komponen parametrik dan nonparametrik uji linieritas lebih baik dibandingkan *scatter plot*. Model regresi semiparametrik birespon dengan pendekatan Deret Fourier ini juga bukan model yang sederhana jika dibandingkan dengan model yang diperoleh untuk parameter osilasi yang lebih kecil. Dalam hal ini model belum mampu mencapai K optimal dalam *parsimony model* pada nilai GCV yang minimum. Perlu dilakukan penelitian lanjutan, misalnya dengan melakukan *smoothing* pada model sehingga terdapat parameter penghalus yang dapat mengontrol nilai K optimal pada nilai GCV yang minimum.

BAB 5
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang diperoleh berdasarkan analisis dan pembahasan sebelumnya adalah sebagai berikut :

1. Diberikan model regresi semiparametrik birespon dengan variabel komponen parametrik $x_1, x_2 \ldots, x_p$, dan variabel komponen nonparametrik $t_1, t_2 \ldots, t_R$ mengikuti model :

$$y_{ji} = \beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi} + \sum_{r=1}^{R} f_j(t_{ri}) + \varepsilon_{ji}$$

dimana $i = 1,2,\ldots,n$ dan $j = 1,2$. Komponen nonparametrik didekati dengan fungsi Deret Fourier berbentuk :

$$f_j(t_{ri}) = \gamma_{jr} t_{ri} + \frac{\alpha_{0jr}}{2} + \sum_{k=1}^{K} (\alpha_{kjr} \cos k t_{ri})$$

Parameter K merupakan parameter osilasi. Dengan optimasi WLS (*Weighted Least Square*) :

$$\min_{\boldsymbol{\beta} \in R^{2(p+1)}, \boldsymbol{\eta} \in R^{2(r(K+1)+1)}} \{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta})^{\mathrm{T}} \mathbf{W} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{T}\boldsymbol{\eta})\}$$

diperoleh estimator untuk komponen parametrik

$$\hat{\boldsymbol{\beta}} = \mathbf{A}(\mathbf{K})\mathbf{y}$$

dengan

$$\mathbf{A}(\mathbf{K}) = \mathbf{M}\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\right)^{-1}\left\{\mathbf{X}^{\mathrm{T}} - \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{W}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\right\}\mathbf{W}$$

dan estimator untuk komponen nonparametrik

$$\hat{\boldsymbol{\eta}} = \mathbf{B}(\mathbf{K})\mathbf{y}$$

dengan

$$\mathbf{B}(\mathbf{K}) = \mathbf{N}\left(\mathbf{T}^{\mathrm{T}}\mathbf{W}\mathbf{T}\right)^{-1}\left\{\mathbf{T}^{\mathrm{T}} - \mathbf{T}^{\mathrm{T}}\mathbf{W}\mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\right\}\mathbf{W}$$

Estimator model regresi semiparametrik birespon dengan pendekatan Deret Fourier diberikan oleh:

$$\hat{\mathbf{y}} = \mathbf{C}(\mathbf{K})\mathbf{y}$$

dengan

$$\mathbf{C}(\mathbf{K}) = \mathbf{X}\mathbf{A}(\mathbf{K}) + \mathbf{T}\mathbf{B}(\mathbf{K})$$

2. Dengan menggunakan metode GCV minimum diperoleh model regresi semiparametrik birespon dengan pendekatan Deret Fourier sebagai berikut :

$$\begin{aligned}\hat{y}_{1i} &= 0{,}043 + 0{,}326x_{1i} + 0{,}354x_{2i} + 0{,}182x_{3i} + 0{,}337x_{4i} - 0{,}042t_{1i} + \\ &\quad 0{,}053\cos t_{1i} - 0{,}119\cos 2t_{1i} - 0{,}146\cos 3t_{1i} + 0{,}08t_{2i} + 0{,}035\cos t_{2i} + \\ &\quad 0{,}035\cos 2t_{2i} + 0{,}035\cos 3t_{2i}\end{aligned}$$

$$\begin{aligned}\hat{y}_{2i} &= -1{,}015 + 0{,}939x_{1i} + 0{,}373x_{2i} + 0{,}842x_{3i} - 0{,}31x_{4i} - 0{,}005t_{1i} + \\ &\quad 0{,}226\cos t_{1i} - 0{,}636\cos 2t_{1i} - 1{,}314\cos 3t_{1i} + 0{,}525t_{2i} - 1{,}029\cos t_{2i} + \\ &\quad -1{,}029\cos 2t_{2i} - 1{,}029\cos 3t_{2i}\end{aligned}$$

dengan nilai GCV = 753.586 dan MSE = 3,052 untuk $K = 3$. Model ini memiliki R^2 = 0,9198. Artinya keragaman nilai respon telah mampu dijelaskan oleh variabel prediktor sebesar 91,98%. Model regresi semiparametrik birespon dengan pendekatan Deret Fourier ini dapat digunakan sebagai alternatif dari model regresi parametrik birespon yang melanggar asumsi klasik residual identik, independen dan berdistribusi normal.

5.2 Saran

Berdasarkan analisis dan pembahasan sebelumnya, saran yang disampaikan untuk penelitian selanjutnya adalah sebagai berikut :

1. Perlu dilakukan penelitian lebih lanjut dengan melakukan *smoothing* pada model sehingga terdapat parameter penghalus yang dapat mengontrol nilai K optimal pada nilai GCV minimum. *Smoothing* dilakukan dengan menggunakan *Penalized Least Square* (PLS).
2. Perlu dilakukan penelitian lebih lanjut bagaimana model regresi semiparametrik birespon dengan pendekatan Deret Fourier untuk K yang lebih besar, karena K yang diteliti dalam penelitian ini terbatas.
3. Perlu dilakukan aplikasi jika banyak variabel komponen parametrik dan nonparametrik dalam tiap respon berbeda.
4. Perlu dilakukan penelitian lebih lanjut untuk mendapatkan estimator regresi semiparametrik dengan pendekatan Deret Fourier dalam kasus multirespon beserta aplikasinya.

5. Perlu dilakukan penelitian lebih lanjut mengenai IMH atau *Physical Quality of Life Index* (PQLI) di Indonesia mengingat penggunaan IMH sebagai indeks kesejahteraan sebenarnya juga penting selain mengacu pada IPM.
6. Perlu digali lebih lanjut variabel-variabel lain yang berpengaruh terhadap IPM dan IMH khususnya di Jawa Timur.

“halaman ini sengaja dikosongkan”

DAFTAR PUSTAKA

Antoniadis, A., Gregorie, G., dan Mackeagu, W. (1994), "Wavelet Methods for Curve Estimation", *Journal of the American Statistical Association*, Vol. 89, hal. 1340-1353.

Antoniadis, A., dan Bigot, J., dan Spatinas, T. (2001), "Wavelet Estimators in Nonparametric Regression : A Comparative Simulation Study", *Journal of Statistical Software*, Vol. 6, hal. 1-83.

Asrini, L.J. (2012), "Regresi Semiparametrik Deret Fourier", *Prosiding Seminar Nasional FMIPA Universitas Negeri Surabaya*, hal. 77-80.

Asrini, L.J. (2014), "Fourier Series Semiparametric Regression Models (Case Study: The Production of Law Land Rice Irrigation in Central Java)". *ARPN Journal of Engineering and Applied Sciences*, Vol. 9, hal. 1501-1506.

Badan Pusat Statistik. (2008), *Booklet Indeks Pembangunan Manusia 2006-2007*, BPS, Jakarta.

Badan Pusat Statistik. (2013), *Indeks Pembangunan Manusia*, http://bps.go.id/menutab. php?tabel=1&kat=1&id_subyek=26.

Biedermann, S., Dette, H., Hoffmann, P. (2009), "Constrained Optimal Discrimination Designs for Fourier Regression Models", *Ann Inst Stat Math Journal*, Vol. 61, hal. 143-157.

Bilodeau, M. (1992), "Fourier Smoother and Additive Models", *The Canadian Journal of Statistics*, Vol. 3, hal. 257-259.

Budiantara, I.N., Subanar, dan Soejoeti, Z. (1997), "Weighted Spline Estimator", *Buletin of the International Statistical Institute*, Vol. 51, hal. 333-334.

Budiantara, I.N. (2001), *Regresi Nonparametrik dan Semiparametrik serta Perkembangannya*, Makalah Pembicara Utama pada Seminar Nasional Alumni Pasca Sarjana Matematika Universitas Gadjah Mada, Yogyakarta.

Budiantara, I.N. (2002), "*Estimator Tipe Penalized Likelihood*", *Jurnal Natural FMIPA Unibraw*, Edisi Khusus, hal. 231-235.

Budiantara, I.N., Lestari, B., dan Islamiyati, A. (2010), *Estimator Spline Terbobot dalam Regresi Nonparametrik dan Semiparametrik Heteroskedastik untuk Data Longitudinal*, Hibah Penelitian Kompetensi DP2M Dikti, Jakarta.

Drapper, N.R., dan Smith, H. (1992), *Applied Regression Analysis 2nd Edition*, Marcel Dekker, New York.

Eubank, R.L. (1988), *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York.

Eubank, R.L. (1999), *Spline Smoothing and Nonparametric Regression 2nd Edition*, Marcel Dekker, New York.

Green, P.J., dan Silverman, B.W. (1994), *Nonparametric Regression and Generalized Linear Model,* Chapman and Hall, London.

Haq, M.U. (1996), *Reflections on Human Development 1st Edition,* Oxford University Press, New York.

Hardle, G. (1990), *Applied Nonparametric Regression, Cambridge University Press*, New York.

Hardle, W., Liang, H., dan Gao, J. (2000), *Partially Linear Model*, Physica Verlag, Heidelberg, Berlin.

Islamiyati, A., dan Budiantara, I.N. (2007), "Model Spline dengan Titik – Titik Knot dalam Regresi Nonparametrik", *Jurnal Inferensi*, Vol. 3, hal. 11-21.

Johnson, R.A., dan Wichern, D. (2007). *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey.

Morris, M.D. (1979), *Measuring the Condition of the World's Poor: The Physical Quality of Life Index,* Frank Cass, London.

Morrison, D.F. (2005), *Multivariate Statistical Methods, Fourth Edition*, The Wharton School University of Pennsylvania, Pennsylvania.

Nosal, R., dan Schultz, T. (2008), "PQLI Definition of Critically", *Journal of Pharmaceutical Innovation*, Vol. 3, hal. 69-78.

Pane, R., Budiantara, I.N., Zain, I., dan Otok, B.W. (2014), "Parametric and Nonparametric Estimators in Fourier Series Semiparametric Regression and Their Characteristics". *Applied Mathematical Sciences*, Vol.102, No.8, hal. 5053-5064.

Petrovic, P. (2012), *SPSS Tutorial and Exercise Book for Business Statistics*, Uneversity of Miskolc, Miskolc.

Prahutama, A. (2013), "Model Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur", *Prosiding Seminar Nasional Statistika Undip*, Vol. 10, hal. 69-76.

Rencher, A.C. (2002), *Methods of Multivariate Analysis Second Edition*, John Wiley and Sons Inc., New York.

Rencher, A.C., dan Schaalje, G.B. (2008), *Linear Models in Statistics 2nd Edition*, John Willey and Sons Inc., New Jersey.

Salam, N. (2013), "Estimasi Likelihood Maximum Penalized dari Model Regresi Semiparametrik", *Prosiding Seminar Nasional Statistika Undip*, Vol. 10, hal. 571-582.

Searle, S.R. (1971), *Linear Models*, John Willey and Sons Inc., New York.

Semiati, R. (2010), *Regresi Nonparametrik Deret Fourier Birespon*, Tesis, ITS, Surabaya.

Tripena, A., dan Budiantara, I.N. (2006), *Fourier Estimator in Nonparametric Regression*, International Conference On Natural Sciences and Applied Natural Scienes, Ahmad Dahlan University, Yogyakarta.

Tripena, A. (2007), *Estimator Deret Fourier dalam Regresi Nonparametrik*, Tesis, ITS, Surabaya.

Tripena, A. (2013), "Estimator Deret Fourier untuk Estimasi Kurva Regresi Nonparametrik Birespon", *Magistra*, Vol. 84, hal. 1-15.

United Nations Development Program. (1995), *Human Development Report 1995*, Oxford University Press for UNDP, New York.

United Nations Development Program. (2013), *Human Development Report 2013 The Rise of the South: Human Progress in a Diverse World,* UNDP, New York.

Wahba, G. (1990), *Spline Model for Observational Data*, *SIAM*, XII, Philadelphia.

"halaman ini sengaja dikosongkan"

LAMPIRAN

Lampiran 1. Data Penelitian Prediktor

i	Wilayah	Indeks Kesehatan (x1)	Indeks Pendidikan (x2)	PDRB (x3)	Indeks PPP (x4)	TPAK (t1)	Persentase Kemiskinan (t2)
1	Pacitan	77,82	76,56	6,73	64,26	79,73	17,22
2	Ponorogo	75,67	75,28	6,52	64,78	73,41	11,72
3	Trenggalek	78,55	78,16	6,62	65,57	77,32	14,15
4	Tulungagung	78,25	80,72	6,99	64,37	72,21	9,37
5	Blitar	77,17	77,82	6,35	68,3	73,61	10,7
6	Kediri	75,25	79,05	6,98	63,85	69,86	13,66
7	Malang	74,17	76,22	7,44	65,45	70,26	11
8	Lumajang	71,25	72,01	6,43	63,73	67,51	12,36
9	Jember	63,68	70,86	7,21	63,43	64,13	11,76
10	Banyuwangi	72,30	74,84	7,21	64,47	73,37	9,93
11	Bondowoso	64,75	67,03	6,45	63,17	70,53	15,75
12	Situbondo	64,20	66,02	6,54	64,97	69,37	14,29
13	Probolinggo	61,17	66,80	6,55	65,09	75,31	22,14
14	Pasuruan	66,02	75,95	7,23	65,54	70,4	11,53
15	Sidoarjo	76,72	87,24	7,13	68,13	66,7	6,42
16	Mojokerto	76,07	80,43	7,21	66,76	70,13	10,67
17	Jombang	75,47	80,45	6,97	65,65	66,54	12,18
18	Nganjuk	73,88	77,64	6,68	64,34	67,52	13,17
19	Madiun	73,75	76,28	6,43	62,6	69,99	13,65
20	Magetan	77,77	78,15	6,39	65,62	72,02	11,45
21	Ngawi	75,95	72,64	6,58	62,02	65,5	15,94
22	Bojonegoro	70,70	71,51	5,68	61,02	69,41	16,6
23	Tuban	72,02	71,75	6,22	63,77	66,55	17,77
24	Lamongan	72,58	76,05	7,13	64,51	68,29	16,64
25	Gresik	71,70	80,61	7,43	66,4	63,49	14,29
26	Bangkalan	64,42	68,02	6,31	64,63	70,25	24,61
27	Sampang	64,97	55,46	6,12	64,57	76,69	27,87
28	Pamekasan	66,32	70,18	6,32	63,03	77,48	19,53
29	Sumenep	66,78	65,15	6,33	67,3	76,84	21,87
30	Kota Kediri	76,43	87,82	7,51	67,3	66,93	8,11
31	Kota Blitar	79,67	86,60	6,78	68,66	64,56	6,72
32	Kota Malang	76,70	89,72	7,57	68,87	64,26	5,19
33	Kota Probolinggo	76,43	80,98	6,89	68,9	67,65	18,33
34	Kota Pasuruan	69,10	84,83	6,46	69,06	67,97	7,87
35	Kota Mojokerto	78,33	87,24	7,11	68,45	71,04	6,46
36	Kota Madiun	77,37	88,47	7,79	66,66	62,53	5,35
37	Kota Surabaya	77,55	88,02	7,62	69,42	66,12	6,23
38	Kota Batu	75,00	84,52	8,25	66,73	70,09	4,45

Lampiran 2. Data Penelitian Respon

i	Wilayah	IPM (Y1)	IMH (Y2)
1	Pacitan	72,88	87,13
2	Ponorogo	71,91	84,18
3	Trenggalek	74,09	88,2
4	Tulungagung	74,45	88,47
5	Blitar	74,43	86,7
6	Kediri	72,72	85,09
7	Malang	71,94	83,23
8	Lumajang	69	78,72
9	Jember	65,99	69,82
10	Banyuwangi	70,53	80,43
11	Bondowoso	64,98	69,92
12	Situbondo	65,06	68,61
13	Probolinggo	64,35	65,91
14	Pasuruan	69,17	74,68
15	Sidoarjo	77,36	88,25
16	Mojokerto	74,42	86,43
17	Jombang	73,86	85,58
18	Nganjuk	71,96	83,07
19	Madiun	70,88	82,48

i	Wilayah	IPM (Y1)	IMH (Y2)
20	Magetan	73,85	86,88
21	Ngawi	70,2	83,19
22	Bojonegoro	67,74	77,69
23	Tuban	69,18	79,62
24	Lamongan	71,05	81,04
25	Gresik	75,97	88,65
26	Bangkalan	65,69	70,34
27	Sampang	61,67	66,06
28	Pamekasan	66,51	72,6
29	Sumenep	66,41	71,41
30	Kota Kediri	77,14	87,91
31	Kota Blitar	78,31	80,68
32	Kota Malang	78,43	88,33
33	Kota Probolinggo	75,44	86,17
34	Kota Pasuruan	74,33	80,74
35	Kota Mojokerto	78,01	81,42
36	Kota Madiun	77,5	80,83
37	Kota Surabaya	78,33	89,11
38	Kota Batu	75,42	86,54

Lampiran 3. Data yang Diperlukan untuk Menghitung IMH

i	Wilayah	AHH (Z1)	AKB (Z2)	AMH (Z3)
1	Pacitan	71,69	22,63	91,63
2	Ponorogo	70,40	27,03	88,99
3	Trenggalek	72,13	21,41	92,88
4	Tulungagung	71,95	22,02	94,57
5	Blitar	71,30	23,71	92,05
6	Kediri	70,15	27,79	92,87
7	Malang	69,50	30,46	90,73
8	Lumajang	67,75	37,89	86,58
9	Jember	63,21	56,33	83,65
10	Banyuwangi	68,38	34,81	88,08
11	Bondowoso	63,85	53,93	80,72
12	Situbondo	63,52	54,94	78,31
13	Probolinggo	61,70	63,51	80,48
14	Pasuruan	64,61	51,07	91,17
15	Sidoarjo	71,03	24,27	97,79
16	Mojokerto	70,64	25,54	94,16
17	Jombang	70,28	27,56	93,87
18	Nganjuk	69,33	31,12	91,11
19	Madiun	69,25	31,18	89,61
20	Magetan	71,66	22,85	91,08
21	Ngawi	70,57	27,06	85,58
22	Bojonegoro	67,42	38,67	84,85
23	Tuban	68,21	34,41	85,86
24	Lamongan	68,55	33,72	88,76
25	Gresik	71,47	23,27	97,17
26	Bangkalan	63,65	54,56	82,90
27	Sampang	63,98	54,48	69,12
28	Pamekasan	64,79	50,69	84,21
29	Sumenep	65,07	48,82	78,71
30	Kota Kediri	70,86	24,85	97,60
31	Kota Blitar	72,80	19,50	97,31
32	Kota Malang	71,02	24,74	98,34
33	Kota Probolinggo	70,86	25,12	92,55
34	Kota Pasuruan	66,46	39,45	97,07
35	Kota Mojokerto	72,00	21,88	97,18
36	Kota Madiun	71,42	23,24	97,84
37	Kota Surabaya	71,53	23,18	98,35
38	Kota Batu	70,00	28,87	98,32

Lampiran 4. Program GCV Regresi Semiparametrik Birespon Deret Fourier untuk *p*=2, *r*=4

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
t1=as.matrix(data[,7])
t2=as.matrix(data[,8])
t3=as.matrix(data[,9])
t4=as.matrix(data[,10])
x3=t1
x4=t2
x5=t3
x6=t4
t = cbind(t1,t2,t3,t4)
y = rbind(y1,y2)
p = 2
r = 4
n = 38

fourier = function(y1,y2,x1,x2,t1,t2,t3,t4,p,k,n)
{
x3=t1
x4=t2
x5=t3
x6=t4
t = cbind(t1,t2,t3,t4)
y = rbind(y1,y2)
N<-length(y1)
X<-matrix(1,N,p+1)
X[,2:(p+1)]<-cbind(x1,x2)
X2 = matrix (0,n,p+1)
X1 = cbind (X,X2)
X2 = cbind(X2,X)
X = rbind (X1,X2)

#=====================================================#
bb=r*(k+1)+1
T1 = matrix(1,n,bb)
for (i in 1:r)
{
T1[,i+1]=t[,i]
for (j in 1:k)
{
T1[,i+j+1]=cos(j*t[,i])
}
}

T2 = matrix (0,n,(r*(k+1)+1))
T3 = rbind(T1,T2)
```

Lanjutan Lampiran 4. Program GCV Regresi Semiparametrik Birespon Deret Fourier untuk *p*=2, *r*=4

```
T4 = rbind(T2, T1)
T  = cbind(T3,T4)

#=====================================================#


S1 = matrix(0,n,n)
S2 = matrix(0,n,n)
S3 = matrix(cov(y1,y2),n,n)
S4 = matrix(cov(y1,y2),n,n)

for (i in 1:n)
{
S1[i,i]=var(y1)
S2[i,i]=var(y2)
}


S11 = cbind(S1,S3)
S22 = cbind(S4,S2)

library (MASS)
W= rbind(S11,S22)

#=======================================================#

I1= diag(1,2*(p+1))

M = ginv(I1-
(ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*
%X))

Ak=M%*%(ginv(t(X)%*%W%*%X))%*%(t(X)-
t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T))%*%W

beta=Ak%*%y

I2 = diag(1,2*((r*(k+1))+1))

N = ginv(I2-
(ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*
%T))

Bk=N%*%(ginv(t(T)%*%W%*%T))%*%(t(T)-
t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X))%*%W

nuhat=Bk%*%y
```

Lanjutan Lampiran 4. Program GCV Regresi Semiparametrik Birespon Deret Fourier untuk *p*=2, *r*=4

```
Ck=X%*%Ak+T%*%Bk

yhat = X%*%beta+T%*%nuhat

yhat=Ck%*%y
#==============================================================#
I = diag(1,2*n)
MSE = (1/(2*n))*t(y)%*%t(I-Ck)%*%(I-Ck)%*%y

GCV = MSE/(1/(2*n))*sum(diag(I-Ck))^2
print(GCV)
}
```

Lampiran 5. Program Estimasi Regresi Semiparametrik Birespon Deret Fourier untuk p=2, r=4

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
t1=as.matrix(data[,7])
t2=as.matrix(data[,8])
t3=as.matrix(data[,9])
t4=as.matrix(data[,10])
x3=T1
x4=T2
x5=T3
x6=T4
t = cbind(t1,t2,t3,t4)
y = rbind(y1,y2)
p = 2
r = 4
n = 38

fourier = function(y1,y2,x1,x2,t1,t2,t3,t4,p,k,n)
{
x3=t1
x4=t2
x5=t3
x6=t4
t = cbind(t1,t2,t3,t4)
y = rbind(y1,y2)
N<-length(y1)
X<-matrix(1,N,p+1)
X[,2:(p+1)]<-cbind(x1,x2)
X2 = matrix (0,n,p+1)
X1 = cbind (X,X2)
X2 = cbind(X2,X)
X = rbind (X1,X2)

#======================================================#
bb=r*(k+1)+1
T1 = matrix(1,n,bb)
for (i in 1:r)
{
T1[,i+1]=t[,i]
for (j in 1:k)
{
T1[,i+j+1]=cos(j*t[,i])
}
}

T2 = matrix (0,n,(r*(k+1)+1))
T3 = rbind(T1,T2)
```

Lanjutan Lampiran 5. Program Estimasi Regresi Semiparametrik Birespon Deret Fourier untuk p=2, r=4

```
T4 = rbind(T2, T1)
T  = cbind(T3,T4)

#======================================================#


S1 = matrix(0,n,n)
S2 = matrix(0,n,n)
S3 = matrix(cov(y1,y2),n,n)
S4 = matrix(cov(y1,y2),n,n)

for (i in 1:n)
{
S1[i,i]=var(y1)
S2[i,i]=var(y2)
}


S11 = cbind(S1,S3)
S22 = cbind(S4,S2)

library (MASS)
W= rbind(S11,S22)

#========================================================#

I1= diag(1,2*(p+1))

M = ginv(I1-
(ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*
%X))

Ak=M%*%(ginv(t(X)%*%W%*%X))%*%(t(X)-
t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T))%*%W

beta=Ak%*%y

I2 = diag(1,2*((r*(k+1))+1))

N = ginv(I2-
(ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*
%T))

Bk=N%*%(ginv(t(T)%*%W%*%T))%*%(t(T)-
t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X))%*%W

nuhat=Bk%*%y
```

Lanjutan Lampiran 5. Program Estimasi Regresi Semiparametrik Birespon Deret Fourier untuk *p*=2, *r*=4

```
Ck=X%*%Ak+T%*%Bk

yhat = X%*%beta+T%*%nuhat

yhat=Ck%*%y

#=================================================================#
I = diag(1,2*n)
MSE = (1/(2*n))*t(y)%*%t(I-Ck)%*%(I-Ck)%*%y

GCV = MSE/(1/(2*n))*sum(diag(I-Ck))^2


Rsq = sum((yhat - mean(y))^2)/sum((y - mean(y))^2) #Masalah

cat("GCV = ", GCV,"\n")
cat("MSE = ", MSE,"\n")
cat("beta","\n")
print (beta)
cat("nuhat","\n")
print (nuhat)
cat("yhat","\n")
print (yhat)
}
```

Lampiran 6. Program GCV Regresi Semiparametrik Birespon Deret Fourier untuk *p*=4 *r*=2

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
x3=as.matrix(data[,7])
x4=as.matrix(data[,8])
t1=as.matrix(data[,9])
t2=as.matrix(data[,10])
x5=t1
x6=t2
t = cbind(t1, t2)
y = rbind(y1,y2)
p = 4
r = 2
n = 38

fourier = function(y1,y2,x1,x2,x3,x4,t1,t2,p,k,n)
{
x5=t1
x6=t2
t = cbind(t1,t2)
y = rbind(y1,y2)
N<-length(y1)
X<-matrix(1,N,p+1)
X[,2:(p+1)]<-cbind(x1,x2,x3,x4)
X2 = matrix (0,n,p+1)
X1 = cbind (X,X2)
X2 = cbind(X2,X)
X = rbind (X1,X2)

#=====================================================#
bb=r*(k+1)+1
T1 = matrix(1,n,bb)
for (i in 1:r)
{
T1[,i+1]=t[,i]
for (j in 1:k)
{
T1[,i+j+1]=cos(j*t[,i])
}
}

T2 = matrix (0,n,(r*(k+1)+1))
T3 = rbind(T1,T2)
T4 = rbind(T2, T1)
T  = cbind(T3,T4)

#=====================================================#
```

Lanjutan Lampiran 6. Program GCV Regresi Semiparametrik Birespon Deret Fourier untuk *p*=4 *r*=2

```
S1 = matrix(0,n,n)
S2 = matrix(0,n,n)
S3 = matrix(cov(y1,y2),n,n)
S4 = matrix(cov(y1,y2),n,n)

for (i in 1:n)
{
S1[i,i]=var(y1)
S2[i,i]=var(y2)
}


S11 = cbind(S1,S3)
S22 = cbind(S4,S2)

library (MASS)
W= rbind(S11,S22)

#=============================================================#

I1= diag(1,2*(p+1))

M = ginv(I1-
(ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*
%X))

Ak=M%*%(ginv(t(X)%*%W%*%X))%*%(t(X)-
t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T))%*%W

beta=Ak%*%y

I2 = diag(1,2*((r*(k+1))+1))

N = ginv(I2-
(ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*
%T))

Bk=N%*%(ginv(t(T)%*%W%*%T))%*%(t(T)-
t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X))%*%W

nuhat=Bk%*%y

Ck=X%*%Ak+T%*%Bk

yhat = X%*%beta+T%*%nuhat

yhat=Ck%*%y

#=============================================================#
```

Lanjutan Lampiran 6. Program GCV Regresi Semiparametrik Birespon Deret Fourier untuk *p*=4 *r*=2

```
I = diag(1,2*n)
MSE = (1/(2*n))*t(y)%*%t(I-Ck)%*%(I-Ck)%*%y

GCV = MSE/(1/(2*n))*sum(diag(I-Ck))^2
print(GCV)
}
```

Lampiran 7. Program Estimasi Regresi Semiparametrik Birespon Deret Fourier untuk *p*=4, *r*=2

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
x3=as.matrix(data[,7])
x4=as.matrix(data[,8])
t1=as.matrix(data[,9])
t2=as.matrix(data[,10])
x5=t1
x6=t2
t = cbind(t1, t2)
y = rbind(y1,y2)
p = 4
r = 2
n = 38

fourier = function(y1,y2,x1,x2,x3,x4,t1,t2,p,k,n)
{
x5=t1
x6=t2
t = cbind(t1,t2)
y = rbind(y1,y2)
N<-length(y1)
X<-matrix(1,N,p+1)
X[,2:(p+1)]<-cbind(x1,x2,x3,x4)
X2 = matrix (0,n,p+1)
X1 = cbind (X,X2)
X2 = cbind(X2,X)
X = rbind (X1,X2)

#=====================================================#
bb=r*(k+1)+1
T1 = matrix(1,n,bb)
for (i in 1:r)
{
T1[,i+1]=t[,i]
for (j in 1:k)
{
T1[,i+j+1]=cos(j*t[,i])
}
}

T2 = matrix (0,n,(r*(k+1)+1))
T3 = rbind(T1,T2)
T4 = rbind(T2, T1)
T  = cbind(T3,T4)

#=====================================================#
```

Lanjutan Lampiran 7. Program Estimasi Regresi Semiparametrik Birespon Deret Fourier untuk p=4, r=2

```
S1 = matrix(0,n,n)
S2 = matrix(0,n,n)
S3 = matrix(cov(y1,y2),n,n)
S4 = matrix(cov(y1,y2),n,n)

for (i in 1:n)
{
S1[i,i]=var(y1)
S2[i,i]=var(y2)
}


S11 = cbind(S1,S3)
S22 = cbind(S4,S2)

library (MASS)
W= rbind(S11,S22)

#============================================================#

I1= diag(1,2*(p+1))

M = ginv(I1-
(ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*
%X))

Ak=M%*%(ginv(t(X)%*%W%*%X))%*%(t(X)-
t(X)%*%W%*%T%*%ginv(t(T)%*%W%*%T)%*%t(T))%*%W

beta=Ak%*%y

I2 = diag(1,2*((r*(k+1))+1))

N = ginv(I2-
(ginv(t(T)%*%W%*%T)%*%t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X)%*%W%*
%T))

Bk=N%*%(ginv(t(T)%*%W%*%T))%*%(t(T)-
t(T)%*%W%*%X%*%ginv(t(X)%*%W%*%X)%*%t(X))%*%W

nuhat=Bk%*%y

Ck=X%*%Ak+T%*%Bk

yhat = X%*%beta+T%*%nuhat

yhat=Ck%*%y

#==============================================================#
```

Lanjutan Lampiran 7. Program Estimasi Regresi Semiparametrik Birespon Deret Fourier untuk p=4, r=2

```
I = diag(1,2*n)
MSE = (1/(2*n))*t(y)%*%t(I-Ck)%*%(I-Ck)%*%y

GCV = MSE/(1/(2*n))*sum(diag(I-Ck))^2


Rsq = sum((yhat - mean(y))^2)/sum((y - mean(y))^2)

cat("GCV = ", GCV,"\n")
cat("MSE = ", MSE,"\n")
cat("Rsq = ", Rsq,"\n")
cat("beta","\n")
print (beta)
cat("nuhat","\n")
print (nuhat)
cat("yhat","\n")
print (yhat)
}
```

Lampiran 8. Program Estimasi Regresi Parametrik Birespon

```
clear all;
close all;
clc;

%=====INPUT DATA & STATISTIK DESKRIPTIF=====
%Input Data, data_kesda() adalah function data base di Matlab
[Y1, Y2, X1, X2, X3, X4, X5, X6] = data_kesda();

Z=[X1 X2 X3 X4 X5 X6];
Y=[Y1 Y2];
figure(1)
columnname = {'Y1','Y2'};
uitable('data',Y,'ColumnName',columnname);
figure(2)
columnname = {'X1','X2','X3','X4','X5','X6'};
uitable('data',Z,'ColumnName',columnname);
%Rata-Rata Variabel Respon dan Prediktor
miuY1=mean(Y1)
miuY2=mean(Y2)
miuX1=mean(X1)
miuX2=mean(X2)
miuX3=mean(X3)
miuX4=mean(X4)
miuX5=mean(X5)
miuX6=mean(X6)

%Matriks Korelasi
R=[corr(Y1,Y1) corr(Y1,Y2);
   corr(Y2,Y1) corr(Y2,Y2)]

%=====UJI BARTLETT SPHERICITY VARIABEL RESPON (UJI KEBEBASAN)=====
[n,q]=size(Y);
ChiSquare_Hit=-(n-1-(2*q+5)/6)*log(det(R))
if ChiSquare_Hit <= 3.842
        disp ('antar variabel respon independent')
    else
        disp ('antar variabel respon dependent')
end

%=====UJI VARIABEL RESPON BERDISTRIBUSI NORMAL MULTIVARIAT=====
%Scatter Plot
figure(3)
gplotmatrix(Z,Y)
xlabel('X')
ylabel('Y')
%Matriks Gabungan Array Data/Vektor Data Y1, Y2, dan Y3
Y=[Y1 Y2;]
%Matriks Varian-Covarian Y
VCy=cov(Y)
%Matriks Invers Varian-Covarian Y
```

Lanjutan Lampiran 8. Program Estimasi Regresi Parametrik Birespon

```
VCyi=inv(VCy);
%Vektor Obyek Pengamatan - Vektor Mean
Y1m=Y1-mean(Y1);
Y2m=Y2-mean(Y2);
[m a] = size(Y1)
for i=1:m
    d_kuadrat(i) = [(Y1(i,1)-miuY1) (Y2(i,1) - miuY2)]...
        *VCyi*[(Y1(i,1)-miuY1) (Y2(i,1) - miuY2)]';
end

for i=1:m
    if d_kuadrat(i)< 1.386
        N(i)=d_kuadrat(i);
    end
end
d_kuadrat
N
% Uji Hipotesis Variabel Respon Berdistribusi Normal Multivariat
Banyaknya = 0;
for i=1:m
    if N(i)>0
        Banyaknya = Banyaknya + 1;
    else
        Banyaknya = Banyaknya;
    end
end
Banyaknya
PersenData= (Banyaknya/n)*100 %n=38
if Banyaknya >= 0.5*n
        disp ('variabel respon berdistribusi normal multivariat')
    else
        disp ('variabel respon tidak berdistribusi normal
multivariat')
end

% =====ANALISIS REGRESI MULTIVARIAT=====
O=ones(38,1); %vektor 1
X=[O X1 X2 X3 X4 X5 X6;]
% ESTIMASI PARAMETER
(X.')*(X)
inv((X.')*(X))
% Koefisien Regresi
Beta1=inv((X.')*(X))*(X.'*Y1);
Beta2=inv((X.')*(X))*(X.'*Y2);
Beta=[Beta1 Beta2] %koefisien regresi global
% %Residual
epsilon=Y-X*Beta
%=====PENGUJIAN SIGNIFIKANSI PARAMETER=====
%Tabel Multivariat Anova
I=eye(n);
```

Lampiran 8. Program Estimasi Regresi Parametrik Birespon

```
M=X*inv(X'*X)*X';
M0=X(:,1)*inv(X(:,1)'*X(:,1))*X(:,1)';

SSPT=Y'*(I-M0)*Y;
SST1=Y1'*(I-M0)*Y1; %Sum Square Terkoreksi (SST)
SST2=Y2'*(I-M0)*Y2;

SSPE=Y'*(I-M)*Y;
SSE1=Y1'*(I-M)*Y1 %Sum Square Error (SSE)
SSE2=Y2'*(I-M)*Y2

SSPR=SSPT-SSPE;
SSR1=SST1-SSE1; %Sum Square Perlakuan/Model Terkoreksi (SSP)
SSR2=SST2-SSE2;

df_Total_Multivariat=n-trace(M0);
df_Residual_Multivariat=n-2*trace(M)+trace(M'*M);
df_Model_Multivariat=df_Total_Multivariat-df_Residual_Multivariat;

MSE1=SSE1/df_Residual_Multivariat;
MSE2=SSE2/df_Residual_Multivariat;

MSR1=SSR1/df_Model_Multivariat;
MSR2=SSR2/df_Model_Multivariat;

F_Hitung1=MSR1/MSE1;
F_Hitung2=MSR2/MSE2;

%Pengujian Hipotesa Serentak
lamda=(det(Y'*Y-Beta'*X'*Y)/det(Y'*Y-n*mean(Y)*(mean(Y))'))
if lamda <= 0.4746
        disp ('secara keseluruhan parameter tidak sama dengan nol')
    else
        disp ('secara keseluruhan parameter sama dengan nol')
    end
%UJI INDIVIDU T
[n,p]=size(X);
V_beta1=inv(X'*X)*(SSE1/(n-p-1))
V_beta2=inv(X'*X)*(SSE2/(n-p-1))
g_kk1=diag(V_beta1)
g_kk2=diag(V_beta2)
SE_beta1=sqrt(g_kk1)
SE_beta2=sqrt(g_kk2)
SE_beta=[SE_beta1 SE_beta2]

TGLOBAL=Beta./SE_beta

%=====RESIDUAL DATA=====
%Menampilkan residual
```

```
figure(4)
columnname = {'epsilon1','epsilon2'};
uitable('data', epsilon,'ColumnName',columnname);
vektor_epsilon1=epsilon(:,1);
vektor_epsilon2=epsilon(:,2);

%=====UJI BARTLETT SPHERICITY RESIDUAL (UJI KEBEBASAN)=====
[n,q]=size(epsilon);
R_epsilon=corr(epsilon)
ChiSquare_Hit=-(n-1-(2*q+5)/6)*log(det(R_epsilon))
if ChiSquare_Hit <= 3.842
        disp ('residual data independent')
    else
        disp ('residual data tidak independent')
    end

%=====UJI RESIDUAL BERDISTRIBUSI NORMAL MULTIVARIAT=====
%Rata-Rata Residual
miu_epsilon1=mean(epsilon(:,1))
miu_epsilon2=mean(epsilon(:,2))
%Matriks Varian-Covarian Residual
VCepsilon=cov(epsilon)
%Matriks Invers Varian-Covarian Residual
VCepsiloni=inv(VCepsilon)
%Vektor Obyek Pengamatan - Vektor Mean Residual
epsilon1m=Y1-mean(epsilon(:,1));
epsilon2m=Y2-mean(epsilon(:,2));
%Menghitung di^2
for i=1:n
    d_square(i) = [(epsilon(i,1)-miu_epsilon1)...
        (epsilon(i,2) - miu_epsilon2)]*VCepsiloni*[(epsilon(i,1)...
        -miu_epsilon1) (epsilon(i,2) - miu_epsilon2)]';
end

for i=1:n
    if d_square(i)>1.386
        N(i)=d_square(i);
    end
end
d_square
N
% Uji Hipotesis Residual Berdistribusi Normal Multivariat
Banyaknya = 0;
for i=1:n
    if N(i)>0
        Banyaknya_ = Banyaknya;
    else
        Banyaknya = Banyaknya+ 1;
    end
end
Banyaknya
PersenData= (Banyaknya/n)*100
if Banyaknya >= 0.5*n
```

```
        disp ('residual data berdistribusi normal multivariat')
    else
        disp ('residual data tidak berdistribusi normal
multivariat')
end
```

Langkah penggunaan programnya hanya tinggal running, kemudian muncul output program.

Lampiran 9. Langkah Penggunaan Program GCV untuk *p*=2, *r*=4 beserta Output

1. Simpan data dalam format .txt. Susunan data dimulai dari variabel respon pertama, dan kedua, kemudian diikuti dengan variabel prediktor pertama, kedua, dan seterusnya. Misalkan file disimpan di *drive d*.
2. Aktifkan jendela program R
3. Aktifkan *package* : MASS dengan cara klik *Packages* → *Load Package* → Pilih *Package MASS.*
4. Panggil data di *R Console*

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
t1=as.matrix(data[,7])
t2=as.matrix(data[,8])
t3=as.matrix(data[,9])
t4=as.matrix(data[,10])
x3=t1
x4=t2
x5=t3
x6=t4
t = cbind(t1,t2,t3,t4)
y = rbind(y1,y2)
p = 2
r = 4
n = 38
```

5. *Copy* semua *syntax* dalam *function fourier* di Lampiran 4.
6. *Paste* di *R Console*
7. Panggil fungsi dengan mengetik :
   `fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,k,n)`
   k dapat diganti dengan bilangan 1,2,3, dan seterusnya.
8. Klik *Enter* dan tunggu hingga *running* program selesai. Hasil *running* program akan muncul di *R Console*. Sebagai contoh ketik

`fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,3,n)` lalu enter maka hasilnya

```
[,1]
[1,] 6290550
```

Lampiran 10. Langkah – Langkah Penggunaan Program Estimasi untuk p=2, r=4 beserta Output

1. Simpan data dalam format .txt. Susunan data dimulai dari variabel respon pertama, dan kedua, kemudian diikuti dengan variabel prediktor pertama, kedua, dan seterusnya. Misalkan file disimpan di *drive d*.
2. Aktifkan jendela program R
3. Aktifkan *package* : MASS dengan cara klik *Packages* → *Load Package* → Pilih *Package MASS.*
4. Panggil data di *R Console*

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
t1=as.matrix(data[,7])
t2=as.matrix(data[,8])
t3=as.matrix(data[,9])
t4=as.matrix(data[,10])
x3=T1
x4=T2
x5=T3
x6=T4
t = cbind(t1,t2,t3,t4)
y = rbind(y1,y2)
p = 2
r = 4
n = 38
```

5. *Copy* semua *syntax* dalam *function fourier* di Lampiran 5.
6. *Paste* di *R Console*
7. Panggil fungsi dengan mengetik :
   `fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,k,n)`
   k dapat diganti dengan bilangan 1,2,3, dan seterusnya.
9. Klik *Enter* dan tunggu hingga *running* program selesai. Hasil *running* program akan muncul di *R Console*. Sebagai contoh ketik

`fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,3,n)` lalu enter maka hasilnya

```
GCV =  2927628
MSE =  11.45107
beta
              [,1]
[1,] 4.464424e-03
[2,] 3.256730e-01
[3,] 3.538957e-01
[4,] 6.672804e-07
[5,] 9.391147e-01
[6,] 3.728341e-01
```

Lanjutan Lampiran 10. Langkah – Langkah Penggunaan Program Estimasi untuk p=2, r=4 beserta Output

```
nuhat
               [,1]
 [1,]  6.158270e-03
 [2,]  1.775473e-01
 [3,]  3.349607e-01
 [4,] -4.313465e-02
 [5,]  5.335227e-02
 [6,] -1.197427e-01
 [7,] -1.469852e-01
 [8,]  7.852455e-02
 [9,]  6.158270e-03
[10,]  6.158270e-03
[11,]  6.158270e-03
[12,]  6.158270e-03
[13,]  6.158270e-03
[14,]  6.158270e-03
[15,]  6.158270e-03
[16,]  6.158270e-03
[17,]  6.158270e-03
[18,]  9.865100e-07
[19,]  8.423523e-01
[20,] -3.096818e-01
[21,] -4.660590e-03
[22,]  2.255416e-01
[23,] -6.357515e-01
[24,] -1.313714e+00
[25,]  5.244310e-01
[26,]  9.865100e-07
[27,]  9.865100e-07
[28,]  9.865100e-07
[29,]  9.865100e-07
[30,]  9.865100e-07
[31,]  9.865100e-07
[32,]  9.865100e-07
[33,]  9.865100e-07
[34,]  9.865100e-07
yhat
          [,1]
 [1,] 72.87002
 [2,] 71.53995
 [3,] 74.01814
 [4,] 74.20005
 [5,] 74.35714
 [6,] 72.84254
 [7,] 72.14091
 [8,] 68.80247
 [9,] 66.13420
[10,] 70.36013
[11,] 64.87293
```

```
[12,] 65.22531
[13,] 64.43305
[14,] 69.19207
[15,] 77.28963
[16,] 74.71043
[17,] 73.93809
[18,] 71.96130
[19,] 70.84748
[20,] 73.63050
[21,] 70.40037
[22,] 67.91411
[23,] 69.40221
[24,] 71.62127
[25,] 73.72184
[26,] 65.99225
[27,] 61.72711
[28,] 66.27635
[29,] 66.29265
[30,] 77.67129
[31,] 78.25305
[32,] 78.66658
[33,] 75.80018
[34,] 74.51099
[35,] 77.73263
[36,] 77.76988
[37,] 78.36989
[38,] 75.64748
[39,] 92.34116
[40,] 86.16832
[41,] 92.35140
[42,] 90.12310
[43,] 89.43574
[44,] 88.95087
[45,] 87.52738
[46,] 80.51451
[47,] 74.22724
[48,] 83.63597
[49,] 73.69875
[50,] 74.77604
[51,] 71.25348
[52,] 78.16726
[53,] 89.20562
[54,] 90.58703
[55,] 87.52741
[56,] 85.57108
[57,] 86.40122
[58,] 89.49529
[59,] 87.07604
[60,] 83.51090
```

Lanjutan Lampiran 10. Langkah – Langkah Penggunaan Program Estimasi untuk *p*=2, *r*=4 beserta Output

```
[61,] 83.48183
[62,] 87.23039
[63,] 87.59330
[64,] 75.17332
[65,] 72.27289
[66,] 77.29008
[67,] 74.93479
[68,] 93.19108
[69,] 91.48533
[70,] 91.34757
[71,] 89.04274
[72,] 83.38197
[73,] 90.59832
[74,] 91.98787
[75,] 90.23942
[76,] 90.58021
```

Lampiran 11. Langkah Penggunaan Program GCV untuk *p*=4, *r*=2 beserta Output

1. Simpan data dalam format .txt. Susunan data dimulai dari variabel respon pertama, dan kedua, kemudian diikuti dengan variabel prediktor pertama, kedua, dan seterusnya. Misalkan file disimpan di *drive d*.
2. Aktifkan jendela program R
3. Aktifkan *package* : MASS dengan cara klik *Packages* → *Load Package* → Pilih *Package MASS.*
4. Panggil data di *R Console*

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
x3=as.matrix(data[,7])
x4=as.matrix(data[,8])
t1=as.matrix(data[,9])
t2=as.matrix(data[,10])
x5=T1
x6=T2
t = cbind(T1, T2)
y = rbind(y1,y2)
p = 4
r = 2
n = 38
```

5. *Copy* semua *syntax* dalam *function fourier* di Lampiran 6.
6. *Paste* di *R Console*
7. Panggil fungsi dengan mengetik :
   `fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,k,n)`
   k dapat diganti dengan bilangan 1,2,3, dan seterusnya.
8. Klik *Enter* dan tunggu hingga *running* program selesai. Hasil *running* program akan muncul di *R Console*. Sebagai contoh ketik

`fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,3,n)` lalu enter maka hasilnya

```
[,1]
[1,] 753586
```

Lampiran 12. Langkah – Langkah Penggunaan Program Estimasi untuk p=2, r=4 beserta Output

1. Simpan data dalam format .txt. Susunan data dimulai dari variabel respon pertama, dan kedua, kemudian diikuti dengan variabel prediktor pertama, kedua, dan seterusnya. Misalkan file disimpan di *drive d*.
2. Aktifkan jendela program R
3. Aktifkan *package* : MASS dengan cara klik *Packages* → *Load Package* → Pilih *Package MASS.*
4. Panggil data di *R Console*

```
data<- read.table("d:/datatesis.txt", header=TRUE)
y1=as.matrix(data[,3])
y2=as.matrix(data[,4])
x1=as.matrix(data[,5])
x2=as.matrix(data[,6])
x3=as.matrix(data[,7])
x4=as.matrix(data[,8])
t1=as.matrix(data[,9])
t2=as.matrix(data[,10])
x5=T1
x6=T2
t = cbind(T1, T2)
y = rbind(y1,y2)
p = 4
r = 2
n = 38
```

5. *Copy* semua *syntax* dalam *function fourier* di Lampiran 7.
6. *Paste* di *R Console*
7. Panggil fungsi dengan mengetik :
   `fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,k,n)`
   k dapat diganti dengan bilangan 1,2,3, dan seterusnya.
8. Klik *Enter* dan tunggu hingga *running* program selesai. Hasil *running* program akan muncul di *R Console*. Sebagai contoh ketik

`fourier(y1,y2,x1,x2,x3,x4,t1,t2,p,3,n)` lalu enter maka hasilnya

```
GCV =  753586.4
MSE =  3.051896
Rsq =  0.9198171
beta
                [,1]
 [1,]  0.007973942
 [2,]  0.326168092
 [3,]  0.353775370
 [4,]  0.182433830
 [5,]  0.337028748
 [6,]  0.013956847
 [7,]  0.939114549
```

Lanjutan Lampiran 12. Langkah – Langkah Penggunaan Program Estimasi untuk *p*=2, *r*=4 beserta Output

```
 [8,]  0.372834223
 [9,]  0.842351345
[10,] -0.309682353
muhat
            [,1]
 [1,]  0.03483922
 [2,] -0.04225051
 [3,]  0.05312548
 [4,] -0.11857317
 [5,] -0.14621902
 [6,]  0.07975740
 [7,]  0.03483922
 [8,]  0.03483922
 [9,]  0.03483922
[10,] -1.02886741
[11,] -0.00464672
[12,]  0.22551149
[13,] -0.63572118
[14,] -1.31368852
[15,]  0.52452883
[16,] -1.02886741
[17,] -1.02886741
[18,] -1.02886741
yhat
         [,1]
 [1,] 73.21237
 [2,] 71.87738
 [3,] 74.36128
 [4,] 74.53773
 [5,] 74.70218
 [6,] 73.17527
 [7,] 72.48031
 [8,] 69.13332
 [9,] 66.45878
[10,] 70.69869
[11,] 65.19752
[12,] 65.55449
[13,] 64.76322
[14,] 69.52639
[15,] 77.63468
[16,] 75.05262
[17,] 74.27508
[18,] 72.29317
[19,] 71.17466
[20,] 73.96773
[21,] 70.72383
[22,] 68.23355
[23,] 69.72650
[24,] 71.95444
```

Lanjutan Lampiran 12. Langkah – Langkah Penggunaan Program Estimasi untuk *p*=2, *r*=4 beserta Output

```
[25,] 74.05508
[26,] 66.31967
[27,] 62.05783
[28,] 66.60784
[29,] 66.63068
[30,] 78.01295
[31,] 78.59591
[32,] 79.00898
[33,] 76.14279
[34,] 74.84824
[35,] 78.08278
[36,] 78.10813
[37,] 78.71997
[38,] 75.99491
[39,] 88.24018
[40,] 82.06736
[41,] 88.25046
[42,] 86.02216
[43,] 85.33492
[44,] 84.84976
[45,] 83.42643
[46,] 76.41365
[47,] 70.12616
[48,] 79.53510
[49,] 69.59758
[50,] 70.67502
[51,] 67.15220
[52,] 74.06623
[53,] 85.10493
[54,] 86.48617
[55,] 83.42649
[56,] 81.47007
[57,] 82.30012
[58,] 85.39428
[59,] 82.97482
[60,] 79.40987
[61,] 79.38056
[62,] 83.12934
[63,] 83.49220
[64,] 71.07203
[65,] 68.17151
[66,] 73.18899
[67,] 70.83353
[68,] 89.09024
[69,] 87.38453
[70,] 87.24664
[71,] 84.94160
[72,] 79.28109
[73,] 86.49769
```

Lanjutan Lampiran 12. Langkah – Langkah Penggunaan Program Estimasi untuk p=2, r=4 beserta Output

```
[74,] 87.88693
[75,] 86.13873
[76,] 86.47952
```

BIOGRAFI PENULIS

Penulis yang bernama lengkap M. Fariz Fadillah Mardianto lahir di Madiun tanggal 6 Maret 1991 dari pasangan suami istri, M. Taufan dan Sunarsih. Penulis merupakan anak pertama dari dua bersaudara. Penulis telah menempuh pendidikan formal yaitu di SD Negeri 1 Waru Sidoarjo, SMP Negeri 1 Sidoarjo, SMA Negeri 3 Sidoarjo, dan Jurusan Matematika FMIPA ITS angkatan 2009. Selama di S1 penulis mendalami bidang Matematika yang berhubungan dengan Statistika. Penulis sangat tertarik di bidang tersebut. Oleh sebab itu penulis melanjutkan studi untuk program magister di jurusan Statistika FMIPA ITS melalui program Beasiswa *Fresh Graduate* ITS pada Tahun 2013. Penulis diterima dengan NRP 1313201050. Selama menempuh masa perkuliahan, penulis berperan aktif dalam organisasi dan juga kegiatan pengembangan diri. Organisasi seperti Himatika, Ibnu Muqlah, dan Karang Taruna pernah diikuti. Penulis juga aktif dalam kegiatan karya tulis. Penulis mengikuti kompetisi Karya Ilmiah Remaja (KIR) se-Jawa Timur masing – masing pada tahun 2006 di Univesitas Negeri Surabaya, dan tahun 2007 di Universitas Trunojoyo Bangkalan. Dalam kegiatan tersebut penulis selalu mendapatkan juara II. Dalam PKM (Program Kreativitas Mahasiswa) yang dieslenggarakan DIKTI beberapa penelitian penulis dalam PKM yang pernah didanai, 1 penelitian PKM-GT (Gagasan Tertulis) dan 2 penelitian PKMK (Kewirausahaan). Dalam dunia kerja penulis pernah melakukan kerja praktek di Bank Indonesia Kantor Wilayah IV Surabaya, Astra Daihatsu Motor dan aktif sebagai Asisten Dosen Mata Kuliah Kalkulus I serta Kalkulus II di Jurusan Matematika ITS selama 3 tahun. Mulai tahun 2014 sampai sekarang juga aktif sebagai pengajar tidak tetap di Politeknik SAKTI Surabaya dan Universitas Surabaya. Penulis juga aktif dalam berbagai kepanitiaan kegiatan nasional seperti Konferensi Nasional Matematika 2014, serta pernah mengikuti beberapa pelathian *leadership* yang diselenggarakan oleh PT Astra Internasional. Penulis yang biasa dipanggil Fariz dapat dihubungi melalui alamat email fm.fariz@yahoo.com untuk diskusi lebih lanjut terkait tulisan yang pernah dibuat. Semoga Bermanfaat.